CANCER-ASSOCIATED ISOCITRATE DEHYDROGENASE MUTATION: STRUCTURAL BASIS OF ENZYME INACTIVATION

by

Akunna M. Iheanacho

A dissertation submitted to the Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
January 2014

Abstract

Isocitrate dehydrogenase enzymes catalyze the oxidative decarboxylation of isocitrate to α-ketoglutarate. Somatic point mutations in the *IDH1* gene are observed in glioblastoma multiforme and acute myeloid leukemia, with a hot spot identified at amino acid residue R132. These mutant enzymes display neomorphic activity such that enzymes with this mutation now catalyze the reduction of α-ketoglutarate to 2-hydroxyglutarate. Given their tumorigenic potential, mutant IDH1 enzymes are an attractive drug target for cancer therapies since they possess a cancer specific mutation and produce a metabolite that can be used as biomarker.

The structure of IDH1 R132H- α-ketoglutarate-NADP$^+$ was determined to a resolution of 2.1 Å by x-ray crystallography, and demonstrates that the enzyme is an asymmetric homodimer. To determine the binding site of the small molecules A07 and ZAG within IDH1 R132H, we determined the structure of IDH1 R132H-isocitrate-NADP-A07 with data to 3.0 Å and that of IDH1 R132H-isocitrate-NADP-ZAG to 3.3 Å. In both structures, the small molecules bind at an allosteric site at the dimer interface of the semi-open conformation of IDH1 R132H. The allosteric site locks the enzyme in a conformation that prevents the formation of the pre-transition state necessary for the enzyme to undergo catalysis. This work shows that the dimer interface of IDH1 has a pre-defined pocket that participates in catalysis as it transitions between the open and closed conformations. The temperature factor of the binding site area is higher than that of the other portions of the structure, suggesting that this protein-protein interface participates in the conformational changes required for catalysis.

This work shows that the preformed cavity allows for binding of IDH1 to allosteric inhibitors. Using a combination of high throughput screening, inhibition assays, and structural biology, we identified a protein-protein interface that provides another opportunity for structure-based drug design.

Advisor and first reader: Dr. Sandra B. Gabelli

Second reader: Dr. G. William Wong

Dedication

This work is dedicated to my father, Morris Iheanacho,

who sacrificed his own dreams of obtaining a Ph.D. in order to provide for our family

and in memory of

Lt. Nathan L. Anderson, United States Public Health Service.

Acknowledgements

Though my attempts to push back the frontiers of science have sometimes proven to be frustrating, they have also produced some of the most exhilarating moments of my life. In these endeavors, I have been extremely fortunate to have the invaluable support, insight, direction and encouragement of many phenomenal mentors and colleagues in the profession. Their input has been invaluable to me in preparing this body of work.

First, I would like to thank my advisor, Dr. Sandra Gabelli, for granting me the opportunity to work in her research laboratory. Dr. Gabelli has consistently demonstrated her unwavering faith and belief in my capabilities as a scientist and has always made herself available for discussion, trouble-shooting, advice, and other forms of mentoring. I would also like to extend my sincerest gratitude to her for answering my frantic text messages at midnight and during many weekends when my protein seemed to disappear on the column or with the purification step I was running at the time. Her door was always open, and she consistently encouraged me to have a healthy work-life balance.

Next, I wish to extend my thanks to my thesis committee members, Dr. Fredric Wondisford, Dr. Mario Amzel, and Dr. G. William Wong, for their helpful and constructive suggestions. I would like to further extend my gratitude to Dr. Wong for not only serving as a thesis reader, but for also charting my progress outside of thesis committee meetings. I proffer many thanks to Dr. Mario Bianchet who took the time to assist me throughout my learning process. He provided insights that guided my thinking and substantially improved the finished product. In addition, immeasurable thanks to Dr. Bert Vogelstein, Dr. Oleg Schmidt-Kittler, and other members of the Vogelstein laboratory. I was also able to draw on the expertise of Dr. Amzel and received equally vital assistance from him. My work would not have been possible without the collaboration of the Vogelstein and Amzel laboratories. I sincerely appreciate Dr. Jan Hoh, as well, for his support throughout my graduate career and for his continued faith in my scientific capabilities.

Drs. Kathryn Moore and Sara Divall have been instrumental in teaching me everything I know about mouse husbandry. I thank them both for all their help and guidance. Additionally, I am indebted to Sara for not only assisting me with editing my fellowship applications, but also challenging my thinking and my approach with all of my research projects.

Special thanks to Dr. Krisna Doung-Ly and Dr. Anita Ghosh for their technical and critical assistance with learning to use the AKTA FPLC systems. Krisna also helped me learn the intricacies of insect cell expression. Additionally, I sincerely thank Dr. Dan Leahy, as well as members of his lab, for allowing me to use his tissue culture hoods and laboratory space for all of my insect cell expression. Besides the technical assistance noted above, I received support from my colleagues in the Gabelli and Amzel Laboratories. In particular, Agedi Boto, Simon Messing, Srinivas Aripirala, and Andres Hernandez all provided me with incisive, thought-provoking conversations on many occasions.

I thank the members of my extended family, close friends, and my community of faith who have proven to be very supportive in all my endeavors. In particular, I would like to acknowledge the Howells, Anyatownwus, Andersons, Ottleys, Starlings, Dr. Charles Onwulata, and Lauren Griggs. I especially want to thank Jeuelle Sam, Nicole and Nichelle Anderson, as well as Sharlene and Anwar Ottley for providing me with a home away from home. I really enjoyed our potlucks, movie nights, road trips, and outings. I send a special word of thanks to Melissa Andrews and Carmen Starling for contacting me on a weekly, sometimes daily, basis simply to make sure that I was well and successfully moving toward my goals. Their uplifting words of encouragement have helped to spur me on and often provided just the motivation I needed to move beyond an impasse.

I owe my parents, Morris and Shirley Iheanacho, heartfelt thanks for their support over the years. Words cannot adequately express the depth of my gratitude. To my sisters, Chioma and Ngozi; my nephews, Nikolas and Timothy; and my brother-in-law, Tim, my deepest thanks for all of your love and support throughout the years. Chioma, I especially thank you for making yourself available to me any time of day or night. Each of you encouraged me to keep pressing toward my goal. To the love of my life, thank you for enduring all of those 16-hour days, 6-day work weeks, and canceled dates. You taught me the secret of true contentment. Finally, but most importantly, I thank God without whose constancy this milestone would have been impossible.

Table of Contents

List of Tables

List of Figures

1. Introduction

1.1 β-Decarboxylating dehydrogenase family of enzymes

Isocitrate dehydrogenases (IDHs) belong to the β-decarboxylating dehydrogenase family of enzymes that catalyze pyridine dependent-dehydrogenation at the α-carbon followed by metal ion-dependent decarboxylation at the β-carbon of 2-hydroxyacids (Figure 1). Divalent cations, as well as the pyridine nucleotide cofactors, NAD^+ or $NADP^+$, are essential for catalysis. The metal ions, Mg^{2+} or Mn^{2+} stabilize the formation of the substrate complex. The cofactors are similar, the only exception being the phosphate group at the 2' position of the adenylate ribose in $NADP^+$; they serve as

$$\text{2-Hydroxy acid} + NAD(P)^+ \overset{Mg^{2+}}{\rightleftharpoons} \text{2-Oxalo acid} + NAD(P)H + H^+$$

$$\text{2-Oxalo acid} + H^+ \overset{Mg^{2+}}{\rightleftharpoons} \text{2-Keto acid} + CO_2$$

Figure 1. Enzymatic reaction catalyzed by β-decarboxylating dehydrogenases. Each member of this family of enzymes catalyzes an equivalent chemical reaction: dehydrogenation at the α-carbon and subsequent decarboxylation at the β-carbon.

reducing equivalents for biosynthetic reactions and signal transduction pathways within the cell. Other members of this family include: isopropylmalate dehydrogenase, tartrate dehydrogenase, and homoisocitrate dehydrogenase (Figure 2) [1, 2]. These enzymes share a common protein fold at the nucleotide-binding site, which differs slightly from the typical βαβαβ motif observed in most oxidoreductases [1-4].

Isopropylmalate dehydrogenase, an NADP$^+$-dependent dehydrogenase is essential for leucine biosynthesis and is expressed in bacteria, fungi, and plants [4]. It converts isopropylmalate to α-ketoisocaproate. Tartrate dehydrogenase plays a significant role in dicarboxylate and glyoxylate metabolism; it enables bacteria and plants to grow on acetate [2]. It produces oxaloglycolate and NADH from tartrate and NAD$^+$, and is present in pseudomonads [2, 5]. Homoisocitrate dehydrogenase, another NAD$^+$-dependent enzyme, catalyzes the conversion of homoisocitrate to α-ketoadipate, and is crucial for lysine biosynthesis in several different organisms [6].

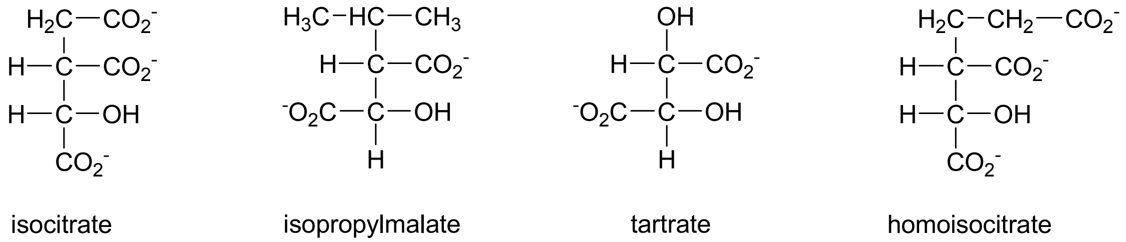


Figure 2. β-decarboxylating dehydrogenase substrates have similar structures. The substrates of these enzymes are structurally similar. Isocitrate, isopropylmalate, tartrate, and homoisocitrate all possess a common malate core [1].

IDHs catalyze the oxidative decarboxylation of isocitrate to α-ketoglutarate (αKG) and are present in many species. NADP$^+$-dependent IDH is ubiquitously expressed in both prokaryotic and eukaryotic organisms and it produces NADPH and α-ketoglutarate (i.e. 2-oxoglutarate) for processes within the cell [7]. It supplies NADPH for the pentose phosphate pathway and is present in the cytosol and the peroxisomes of plants, yeast, mice, rats, and humans [8-11]. Only eukaryotic organisms utilize NAD$^+$-dependent IDH and it is present in the mitochondria. The reaction catalyzed by this enzyme is essentially irreversible and supplies NADH used for ATP production

1.2 Enzyme regulation of IDH

The method of regulating IDH enzyme activity differs among organisms. A single bifunctional protein, IDH Kinase/Phosphatase, uses cycles of protein phosphorylation and dephosphorylation to mediate the activity of prokaryotic NADP$^+$-dependent IDH [12, 13]. In *E. coli,* S113, is present in the active site of IDH, and undergoes phosphorylation to inactivate the enzyme [14]. The addition of the phosphate group prevents substrate binding through steric hindrance and electrostatic repulsion. The removal of the phosphate group by the IDH Kinase/Phosphatase restores enzymatic activity. AMP levels control the entire IDH phosphorylation cycle, with AMP activating IDH phosphatase function and inhibiting the IDH kinase function of this enzyme [15].

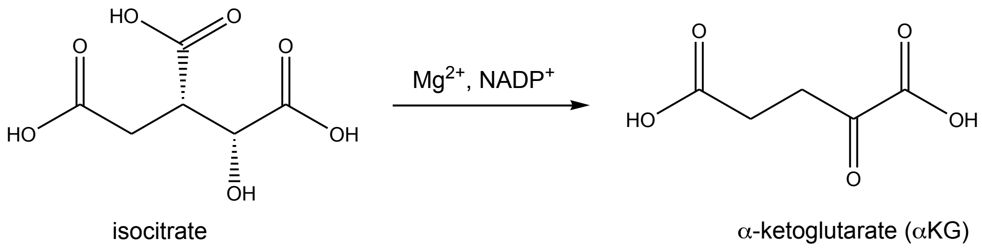


Figure 3. Enzymatic reaction catalyzed by Isocitrate dehydrogenase.
IDH catalyzes the Mg^{2+}-dependent and NADP$^+$-dependent dehydrogenation of ICT to oxalosuccinate and NADPH. Mg^{2+}-dependent decarboxylation of oxalosuccinate to αKG and the loss of CO_2 follow.

In eukaryotic systems, the affinity of IDH for isocitrate is controlled by allosteric regulation. In yeast, AMP allosterically activates IDH, while ATP is restrictive. In mammalian cells, ADP enhances the affinity of IDH for substrate, while ATP is inhibitory [16, 17]. This prevents IDH from catalyzing the reaction unless levels of AMP or ADP are low. Both enzymes are inhibited by NADH and flux through the TCA cycle is diminished when cellular ratios of [NADH]:[NAD^+] and [ATP]:[AMP or ADP] are high. Xu posits that in humans, cytosolic $NADP^+$-dependent IDH activity is controlled by a self-regulatory mechanism in which secondary structural elements protrude into the active site and prevent substrate binding [18]. No other mechanisms of regulation have been described for other IDH enzymes.

1.3 Subcellular localization and cofactor requirement of Human isoforms of IDH

Isocitrate dehydrogenases can be divided based on their subcellular localization and their requirement for either $NADP^+$ or NAD^+ as cofactors. Three isoforms exist in humans: isocitrate dehydrogenase 1 (IDH1), isocitrate dehydrogenase 2 (IDH2), and isocitrate dehydrogenase 3 (IDH3). There are five genes encoding these enzymes: *IDH1* on chromosome 2q33.3, *IDH2* on chromosome 15q26.1, and *IDH3A* on 15q25.1-q25.2, *IDH3B* on 20p13, and *IDH3G* Xq28 [19-22]. The NADP-dependent homologs IDH1 and IDH2 share considerable sequence similarity (70%) and are homodimers (Figure 4), whereas, IDH3 is hetero-oligomeric, and more dissimilar, with two α subunits, one β subunit, and one γ subunit [23, 24]. Although all three subunits participate in catalysis, ADP binding sites important for allosteric regulation are only found on the β and γ subunits of IDH3 [24].

All three isoforms serve different roles within the cell (Figure 5). IDH1 localizes primarily to the cytosol, but it also contains a peroxisomal targeting sequence that allows for import into peroxisomes [8, 25]. IDH1 is involved in cholesterol synthesis, fatty acid synthesis, and regulates lipid metabolism [25, 26]. It replenishes cytosolic stores of αKG essential for the proper functioning of αKG-dependent dioxygenases, such as prolyl hydrolases, ten-eleven translocation (TET) family of 5-methylcytosine hydroxylases, and histone demethylases (KDM). There are several dioxygenases (Table 1) present within the cell, and they catalyze a number of different reactions [27, 28]. In addition to αKG, they require oxygen and iron as cofactors, and possess enhanced activity in the presence

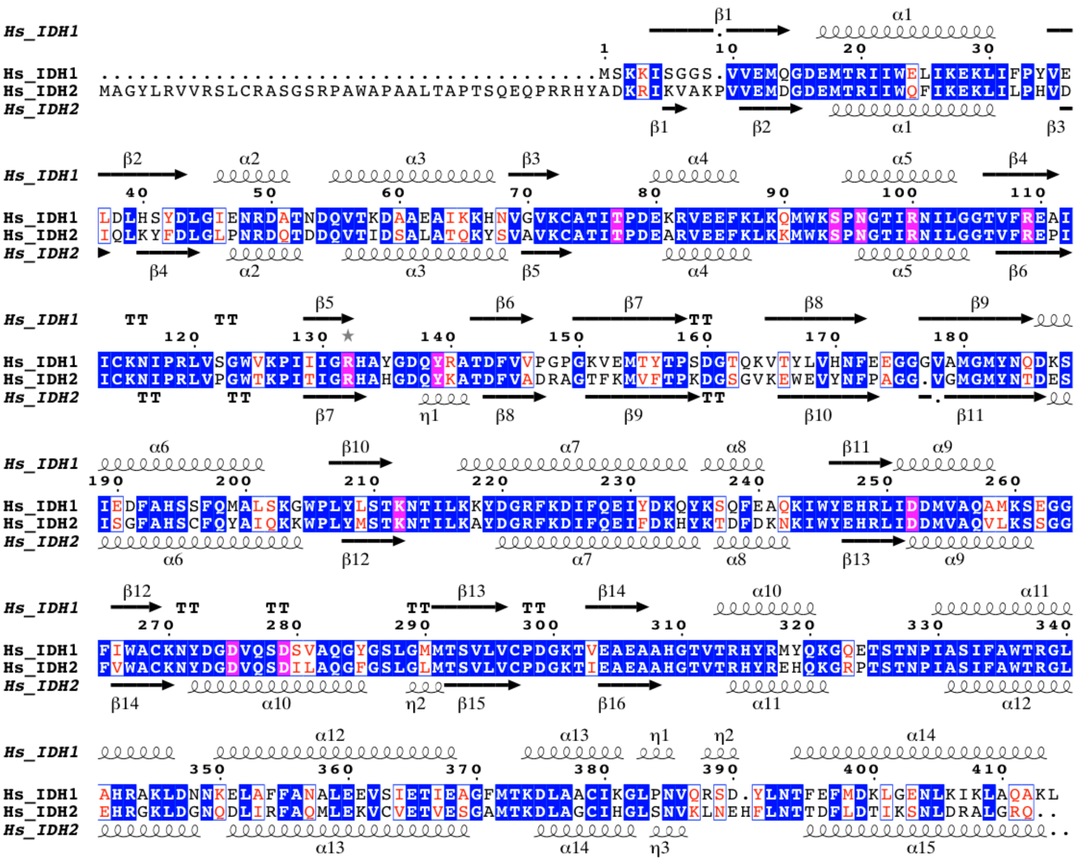


Figure 4. IDH1 and IDH2 share considerable sequence similarity.
Human homologs, IDH1 and IDH2, share 70% sequence similarity but are dissimilar to IDH3. Fully conserved amino acid residues are shaded in blue. Residues important for catalysis are shaded in magenta. The star over amino acid residue 132 of IDH1 and the analogous amino acid residue 172 of IDH2 denote hot spots. The secondary structural assignment above the alignment corresponds to monomer A of *Hs*IDH1, and below the alignment corresponds to monomer A of *Hs*IDH2.

of ascorbate, and only function under normoxic conditions [29, 30]. Additionally, NADPH is a reducing equivalent that supplies power to many synthetic processes [31].

As such, IDH1 protects against oxidative damage and detoxifies reactive oxygen species (ROS) by supplying NADPH for antioxidant systems that utilize glutathione or thioredoxin [31-35].

Although both IDH2 and IDH3 reside in the mitochondria, IDH3 is thought to make a more significant contribution to cellular metabolism and mitochondrial function by providing reducing equivalents that can be oxidized for ATP production. Historically, IDH3 was thought to catalyze the third step of the TCA cycle and the function of these isoforms was not thought to be redundant. However, early work by Alp suggested the possible involvement of both isoforms in the TCA cycle [36]. More recent studies have re-examined their roles and demonstrate that when IDH3 activity is defective, IDH2 can compensate for this deficiency with the help of a nicotinamide-nucleotide transhydrogenase present in the inner membrane of the mitochondrion. This transhydrogenase has the ability to convert NADPH to NADH [37].

Additionally, a recent study of individuals with retinitis pigmentosa, an inherited degenerative eye disease, revealed homozygous mutations occurring in *IDH3B*. Although this mutation effectively eliminates the function of IDH3, abnormalities typically associated with mitochondrial dysfunction such as cardiac dysrhythmias, or reduced muscle strength, were not observed [38]. Kinetic analysis of IDH2 and IDH3 from affected individuals indicated that while IDH2 enzymatic activity was normal, the catalytic activity of IDH3 was significantly reduced [38]. This data suggests that although IDH3 is crucial for normal function in the retina, IDH2 may catalyze the third step of the TCA cycle in other tissues. It is possible that IDH3 serves a regulatory or accessory role in tissues other than the retina. Since mutations in *IDH3* have not been associated with cancer, we will limit our discussion from this point on to IDH1 and IDH2.

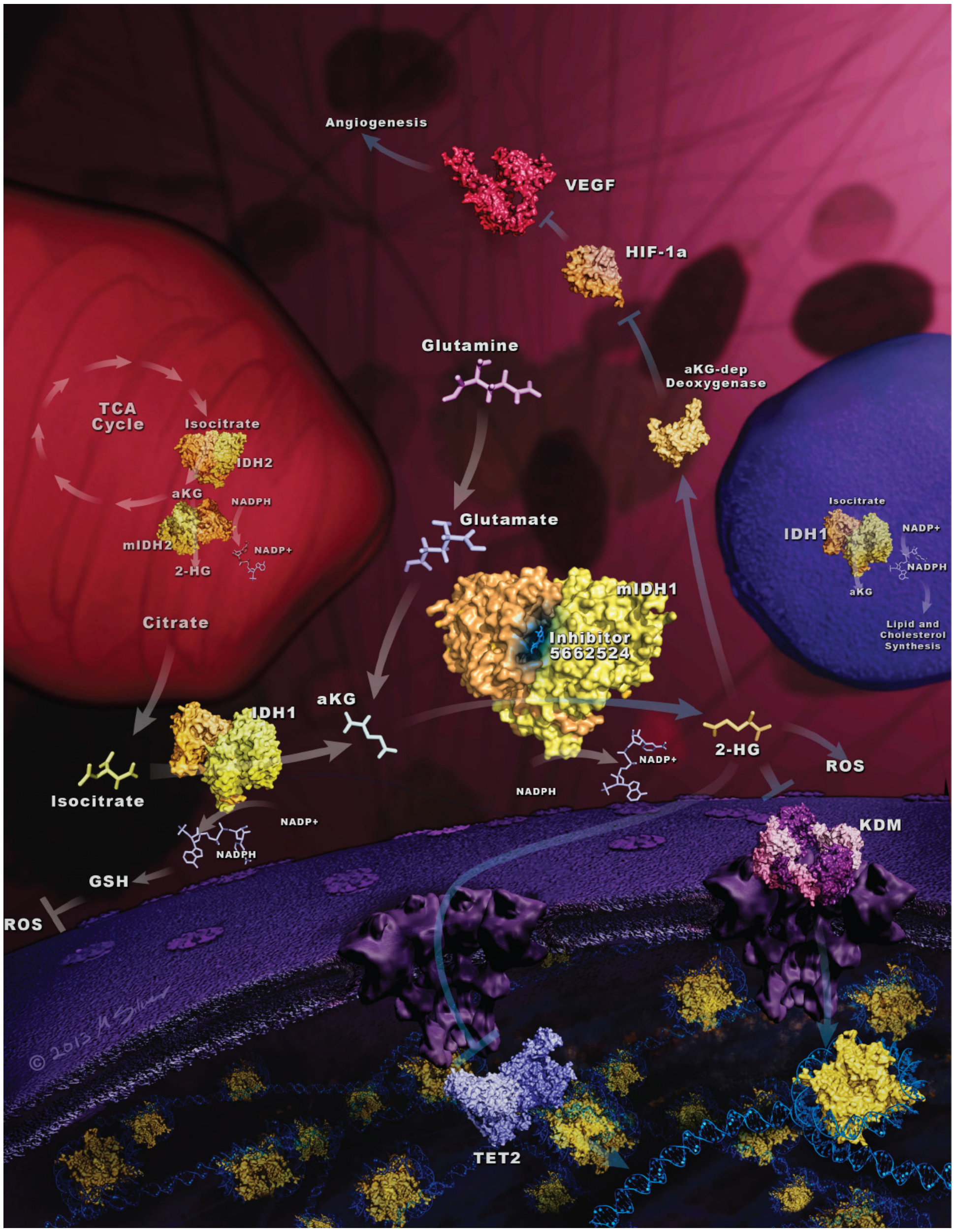


Figure 5. Isocitrate dehydrogenases are involved in many cellular processes.
IDHs serve different roles within the cytosol, peroxisome, and mitochondrion. They participate in cholesterol synthesis, lipid metabolism, and protect against ROS. These enzymes replenish stores of αKG for dioxygenases such as TET2 and KDM. IDHs also provide NADPH as a reducing equivalent for biosynthetic reactions within the cell.

Table 1. Known and proposed roles for human αKG-dependent dioxygenases.

Dioxygenase	Process	Known	Proposed
CPH	Collagen Modification	Stabilize Helix	
P3H	Collagen Modification	Helix Modification	
LH	Collagen Modification	Anchor Point	
TMLH	Carnitine Biosynthesis	FA Transport	
GBBH	Carnitine Biosynthesis	FA Transport	
PhyH	CoA Metabolism	FA Metabolism	
PHD1-3	Hypoxic Sensing	Signal Degradation	
FIH	Hypoxic Sensing	Blocks Transcripton	
FIH	Ankyrin Hydroxylation	HIF Regulation	Protein Degradation
EGFH	GF Hydroxylation	Tumorgenesis	
JMJD	Histone Demethylation	Epigenetic Regulation	
JMJD6	Histone Demethylation	Epigenetic Regulation	
TET2	DNA Demethylation	DNA Repair	Regulation
hABH2/3	DNA Demethylation	DNA Repair	Regulation
FTO	DNA Demethylation	Fat Metabolism	Obesity

Abbreviations include: CPH, collagen proline (4R)-hydoxylase; P3H, proline (3S)-hydroxylase; LH, lysine hydroxylase; TMLH, trimethyl lysine hydroxylase; FA, fatty acid; GBBH, γ-butyrobetaine hydroxylase; PhyH, phytanoyl-CoA hydroxylase; PHD1-3, prolyl hydroxylase domain-containing proteins 1-3; FIH, factor inhibiting HIF; EGFH, epidermal growth factor hydroxylase; GF, growth factor; JMJD, Jumonji domain-containing, hABH2/3, AlkB homologs 2 and 3; TET2, ten-eleven translocase 2; and FTO, fat mass and obesity associated.

Additionally, a recent study of individuals with retinitis pigmentosa, an inherited degenerative eye disease, revealed homozygous mutations occurring in *IDH3B*. Although this mutation effectively eliminates the function of IDH3, abnormalities typically associated with mitochondrial dysfunction such as cardiac dysrhythmias, or reduced muscle strength, were not observed [38]. Kinetic analysis of IDH2 and IDH3 from affected individuals indicated that while IDH2 enzymatic activity was normal, the catalytic activity of IDH3 was significantly reduced [38]. This data suggests that although IDH3 is crucial for normal function in the retina, IDH2 may catalyze the third step of the TCA cycle in other tissues. It is possible that IDH3 serves a regulatory or accessory role in tissues other than the retina. Since mutations in *IDH3* have not been associated with cancer, we will limit our discussion from this point on to IDH1 and IDH2.

1.4 Rossmann fold and common structural features of IDH enzymes

The Rossmann fold is frequently observed in nucleotide-binding proteins and oxidoreductases. It consists of a six-stranded parallel β-sheet surrounded by four α-helices, which form the nucleotide-binding domain of most oxidoreductases [39, 40]. The nucleotide-binding domain contains the conserved signature sequence $GX_{1\text{-}2}GXXG$, in which X is any amino acid residue [41]. This βαβαβ motif is stabilized through hydrogen bonding and van der Waals interactions between glycine residues on the main chain and the pyrophosphate of the nucleotide [41, 42]. Interestingly, there are three oxidoreductases: isopropylmalate dehydrogenase, tartrate dehydrogenase, and a bacterial NAD^+-dependent IDH, which share a protein fold at the nucleotide-binding domain that differs slightly from this typical βαβαβ motif [1, 2, 4].

As with other dimeric $NADP^+$-dependent IDHs, *Hs*IDH possesses similar topology and secondary structure fold with varying sequence identity (Figure 6). In humans, IDH1 forms a homodimer with each monomer consisting of 414 amino acids folded into three domains: large, small, and clasp. Residues 1-103 and 286-414 make up the large domain and contain the Rossmann fold. The small domain consists of amino acid residues 104-136 and 186-285, which form an α/β sandwich. A β-sheet joins the large and small domains together and two clefts flank either side of the β-sheet. The active site cleft is formed by residues from the large and small domains of one subunit, and the small domain of the adjacent subunit. There are two independent active sites present in homodimeric IDHs; and more than likely, a dimer is the functional unit necessary for catalysis. Two stacked, two-stranded anti-parallel β sheets consisting of residues 137-185, form the clasp domain that interlocks the two subunits together, and frames the active site. This domain varies among species in primary sequence, as well as three-dimensional structure of NADP$^+$-dependent IDHs. In *Ec*IDH, the two clasp domains form two anti-parallel α-helices with a four-stranded anti-parallel β-sheet, and in *S. cerevisiae* two stacked layers of four-stranded anti-parallel β-sheets are formed, whereas in mammalian cells, the clasp domains form one layer of four-stranded anti-parallel β-sheets.

1.5 Catalytic mechanism of IDH

Residues from the large and small domains of one subunit, and the small domain of the adjacent subunit constitute each active site. There are binding sites for both ICT and NADP$^+$, but there is no data to support concerted binding, or to suggest that the binding of one influences the binding of the other. Biochemical and structural studies have demonstrated that eukaryotic IDH undergo conformational changes upon substrate binding and release. The states are described as open when IDH is bound to the cofactor NAD(P)$^+$, semi-open for the ternary complex of IDH-NAD(P)$^+$-ICT, and closed for the addition of divalent metal (Mg^{2+}, Ca^{2+}, and Mn^{2+}). Structures of both open and closed forms of IDH have been determined with a productive Michaelis-Menten complex only formed in the closed state. A pseudo-Michaelis-Menten complex has been observed in the presence of product, either ICT for wild type IDH, or αKG in the case of mutant IDH. Closed catalytically active structures have been determined for wild type IDH1 and mutant forms of IDH1 [18, 43].

The proposed reaction mechanism is based on structural studies of *Ss*IDH and *Hs*IDH [18, 44]. There is great sequence similarity (70%) between these two species and many residues are conserved (Figure 6). The reaction is believed to occur in two steps: the oxidation of isocitrate to oxalosuccinate, followed by the loss of a carboxylate group from oxalosuccinate, and protonation of the β-carbon to form α-ketogutarate [45, 46]. The first step involves proton abstraction of the C2 hydroxyl of ICT and transfer of the hydride ion to the C4 of the nicotinamide ring of $NADP^+$ (Figure 7). Several aspartate residues have emerged as possible general base candidates responsible for the initial proton abstraction in studies of *Ec*IDH and *Ss*IDH [3, 47]. Others postulate that active-site water molecules are responsible for the initial proton abstraction, and coordinate with aspartate residues [44]. Arginine residues form a catalytic triad (*Ss*IDH numbering, R101, R110, and R133) that promotes the initial proton abstraction by lowering the pKa of the metal-bound hydroxyl of ICT [24, 48]. The aspartate residues (*Hs*IDH numbering, D252, D275, and D279) coordinate the divalent metal ion.

The transition state undergoes electrostatic stabilization by the magnesium ion; it serves as a counterbalance for the negative charge formed on the hydroxyl oxygen during dehydrogenation. Siebert was the first to propose the oxalosuccinate intermediate and this hypothesis was later confirmed by Laue structure determination [45, 49]. In the second step of the reaction, the β-carboxylate of the oxalosuccinate intermediate is released as CO_2 and replaced by a proton to form the final product, α-ketogutarate. Mutational studies of *Ss*IDH also suggest that T78 or S95 (*Hs*IDH numbering, T77 or S94) assist in generating a good leaving group [50].

Still, the actual mechanism of the final protonation step is unclear. In initial studies of *Ss*IDH, Y140 was thought to function as the general acid that protonates the substrate following decarboxylation of oxalosuccinate [48]. Site-directed mutational studies of *Ss*IDH suggest that a lysine residue (*Ss*IDH numbering, K212) is essential for this step and stabilizes the carbanion formed by the decarboxylation of the substrate [48]. Additionally, the tyrosine is more likely to act as a general base in the reaction. In structural studies of *Ec*IDH, Y160 supports this claim, and seems to be essential for the dehydrogenation step [51]. However, the possibility remains that Y139 (HsIDH numbering) is important for the protonation of β-carbon following the decarboxylation step. Although the reaction mechanism of IDH is well characterized, questions still remain about the initial proton abstraction, and the final step involving proton donation.

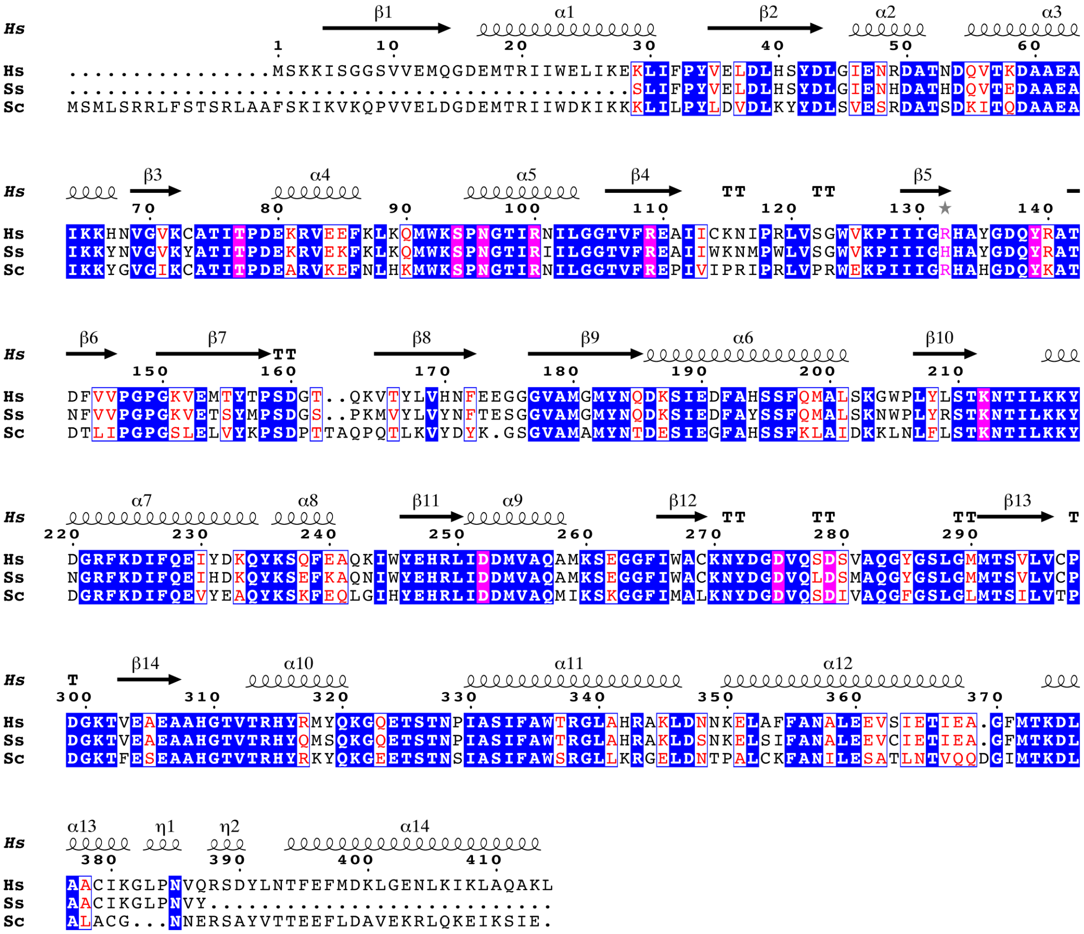


Figure 6. Eukaryotic IDH enzymes possess great structural similarity.
Structure-guided sequence comparison of homodimeric IDH enzymes: *Hs*IDH (human), *Ss*IDH (porcine), and *Sc*IDH (yeast) using *Hs*IDH numbering. Fully conserved amino acid residues are shaded in blue. Residues important for catalysis are shaded in magenta. The secondary structural assignment above the alignment corresponds to monomer A of *Hs*IDH1.

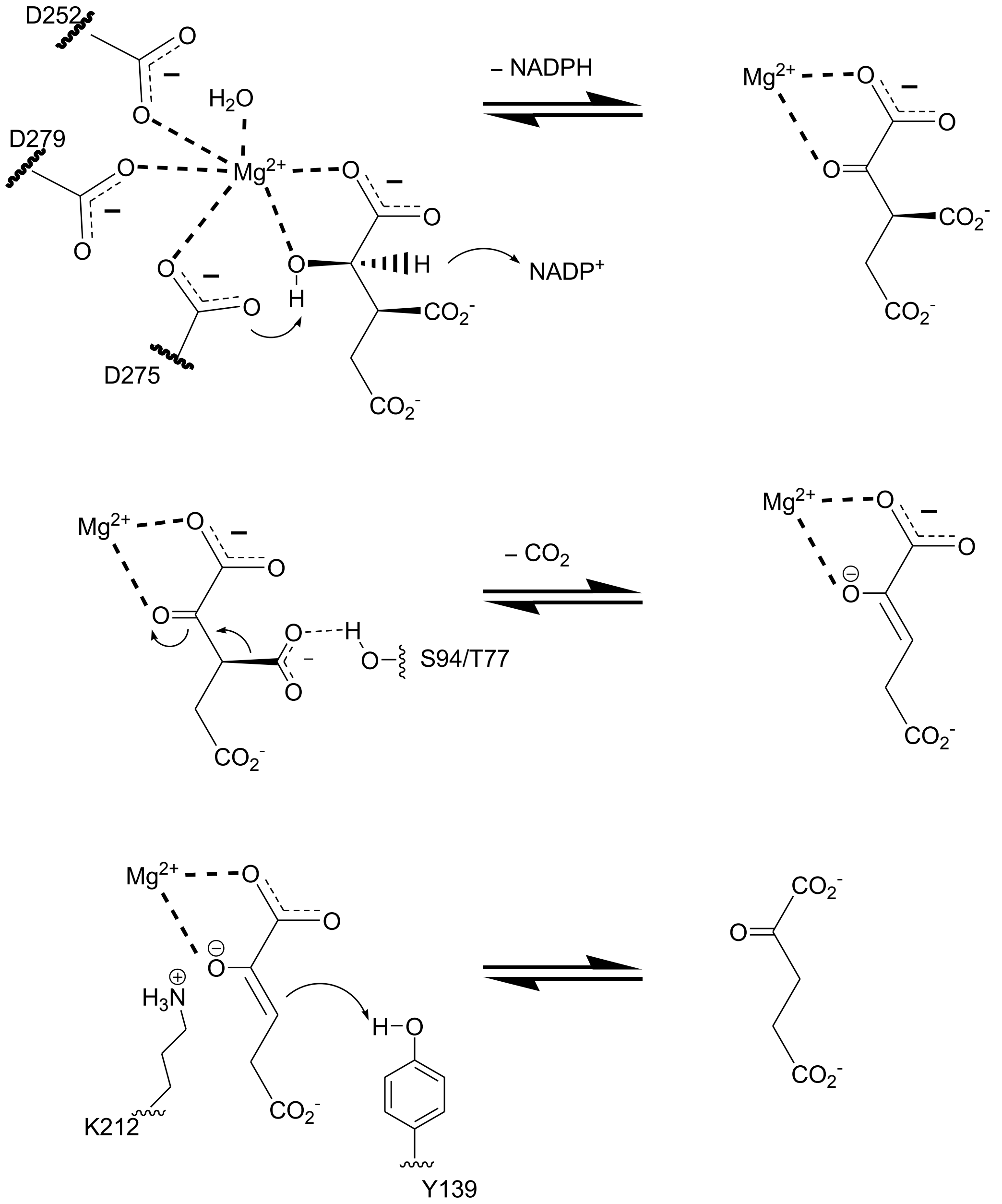


Figure 7. Proposed mechanism for reaction catalyzed by isocitrate dehydrogenase. The reaction begins with proton abstraction followed by hydride transfer to the C4 of the nicotinamide ring of $NADP^+$ to form the intermediate, oxalosuccinate. In the second step, the carboxylate group leaves as CO_2 and a proton replaces it at the β-carbon to form α-ketoglutarate. The transition states are stabilized by Mg^{2+} coordination. *Hs*IDH numbering was utilized for all steps of the proposed mechanism.

2. Links Between Cellular Metabolism and Cancer

2.1 Historical perspective

Since 1927, defects in mitochondrial function have been associated with tumor development and progression. Warburg hypothesized that disruptions in cellular respiration were not only associated with cancer, but were sufficient to drive oncogenesis. He observed that cancer cells converted glucose to lactate at a much higher rate than normal cells even in the presence of oxygen, and this observation became known as the ‘Warburg Effect’ [52-54]. There are four possible fates of glucose once it enters the cell: anaerobically oxidized to pyruvate, shunted into the pentose phosphate pathway, stored as glycogen, or completely oxidized to carbon dioxide and water. Nevertheless, even under normoxic conditions, cancer cells rapidly consume glucose, preferentially converting it to lactate, rather than undergoing oxidative phosphorylation to obtain the maximal amount of ATP [55]. Imaging techniques like positron emission tomography exploit the fact that tumors have increased glucose uptake and radiologists have been able to detect and map tumors with great sensitivity using this technique. Evidence also suggests that more aggressive tumors have a higher glycolytic flux, with poor patient prognosis correlating with higher glucose uptake by tumors in comparison to normal tissues [56, 57].

Aerobic glycolysis is widely accepted as a common feature of cancer cells and has been confirmed in both primary and metastatic tumors. However, it is difficult to provide a definitive explanation for this phenomenon even given this supporting evidence [58, 59]. Although the cell needs energy, it also requires macromolecules. Perhaps the major function of aerobic glycolysis is to provide the cell with glycolytic intermediates for carbon skeletons, energy, and reducing equivalents to facilitate cell division and proliferation. By diverting glucose to other biosynthetic pathways, it enables the cell to generate glycerol for lipogenesis or produce NADPH and ribose-5-phosphate for nucleotide and nucleic acid biosynthesis, while also creating a stockpile of intermediates that can serve as precursors for other pathways.

Historically, genes that cause neoplasia or tumorigenesis are grouped into one of two classes: tumor suppressors or oncogenes. In the case of tumor suppressor genes, inactivating mutations would cause a dominant loss of function that inhibits activity; whereas, activating mutations would cause a gain of function or result in new activity, and are associated with oncogenes. Typically, oncogenes possess recurrent mutations at the same amino acid residues; whereas, tumor suppressor genes are characterized by mutations throughout the entire length of the gene, that alter or truncate the encoded protein [60]. Oncogenes can be activated by mutations, chromosomal rearrangements, or gene amplifications. They encode proteins involved in cell proliferation and apoptosis such as, transcription factors, growth factors, and apoptosis regulators, and can affect the regulation or the structure of the protein encoded by that particular gene.

Based on the classical two-hit hypothesis put forth by Knudson, mutations in tumor suppressor genes were believed to be recessive and affecting both alleles [61]. These genetic disruptions caused their inactivation and led to an overall loss of function. However, over time this view has expanded to include mutated tumor suppressor genes that are haploinsufficient, as well as those that behave in a dominant negative manner.

Haploinsufficiency arises when there are alterations in gene dosage, such that one allele is defective, and the other functional allele is insufficient to sustain normal function. It is difficult to distinguish haploinsufficiency from dominant negative mutations because in both cases the wild-type (WT) allele is retained. However, in the case of dominant negative mutations, mutant gene products recombine with normal WT gene products inactivating them and making non-functional complexes that contain both mutant and WT proteins. Malignant transformation is a multi-step process arising from genetic mutations, epigenetic alterations, cell signaling defects, and other perturbations, that confer a selective growth or survival advantage to the cell [62].

Warburg's theory that mitochondrial dysfunction causes tumor growth and can account for the aerobic glycolysis observed in tumors is still a subject of much debate. Over the years, there has been a shift from viewing oncogenesis as simply a derangement of bioenergetics, to a multifaceted disease caused by the convergence of several factors. Distinct metabolic differences exist between cancer cells and normal cells, and mitochondrial dysfunction has come to the fore once again. Although cellular metabolism is not defined as a hallmark of cancer, recent studies suggest causal links between metabolic alterations and oncogenesis.

2.2 Germline mutations in TCA cycle intermediates

The notion that tumor suppressors and oncogenes can affect the activity and expression of metabolic enzymes has been held for quite some time [59, 63]. Germline mutations in genes that encode tricarboxylic acid (TCA) cycle enzymes have been implicated in several familial cancer cases. The TCA cycle involves a series of enzyme-catalyzed reactions that are essential for aerobic respiration. Succinate dehydrogenease

(SDH), and fumarate hydratase (FH), are sequential enzymes in this cycle. SDH is a heterotetramer that oxidizes succinate to fumarate, and it is also a component of the electron transport machinery, mitochondrial complex II. It resides in mitochondrial membranes and is also responsible for reducing ubiquinone to ubiquinol. Each subunit is encoded by a different gene and heterozygous mutations in *SDHA, SDHB*, *SDHC*, and *SDHD* all result in the complete loss of heterozygosity, (i.e. loss of remaining wild type allele in somatic cells), and render the gene product of *SDH* completely inactive [64-66]. Individuals with mutations in *SDHB* and *SDHC* are afflicted with papillary thyroid carcinomas and renal cell cancers, while mutations in *SDHD* present in gastrointestinal stromal tumors, and colorectal carcinomas. Recent findings indicate that mutations in *SDHA* cause paragangliomas and pheochromocytomas [67], as well as neurodegenerative disease [68].

Fumarate hydratase (FH) is a homotetramer that catalyzes the hydration of fumarate to malate. In humans a single gene, *FH*, encodes two splice variants. The longer form localizes to the mitochondria and participates in TCA cycle, while the shorter form resides in the cytosol and participates in purine synthesis and the urea cycle. Individuals harboring rare heterozygous germline mutations in *FH* developed leiomyomatosis of the skin and uterus, and renal cell carcinomas [69-71]. All mutations described above for fumarate hydratase and succinate dehydrogenase greatly reduce enzymatic activity and lead to a loss of function [72]. Both *FH* and *SDH* have been identified as classical tumor suppressor genes.

The dysfunction of these enzymes and how they drive tumorigenesis is still an area of active study. One possible mechanism by which these mutated genes promote tumorigenesis is by provoking a pseudohypoxic response. Pseudohypoxia is the abberant activation of the hypoxic response. HIF1, hypoxia inducible factor 1, is a transcription factor responsible for mediating the adaptive response to hypoxia and upregulating genes involved in apoptosis, cell survival, and angiogenesis [73-75]. Under hypoxic conditions transcription of these genes is activated and increased expression of HIF1α has been connected to cancer [75, 76]. In the presence of oxygen, dioxygenases will add hydroxyl groups to HIF1α and HIF2α, targeting them for ubiquitination and proteasomal degradation [77]. However, under hypoxic conditions, in the absence of oxygen, these hydroxylation events cannot occur, enabling HIF1α and HIF2α to form dimers with HIF1β and coordinate the hypoxic response. The abnormal accumulation of the intermediates, succinate or fumarate, can cause the inhibition of HIF prolyl hydroxylases, and lead to the stabilization of HIF1α [69]. These initial studies sparked renewed interest in how mitochondrial dysfunction could lead to tumorigenesis. More recent data on mutations affecting another TCA cycle enzyme, IDH, strongly support the notion that defects in cellular metabolism directly contribute to the progression of gliomas.

2.3 Classification of gliomas, common features, and treatment options

There are 69, 720 estimated new cases of malignant brain tumors that are expected to be diagnosed in the United States in 2013. The annual incidence rate of malignant brain tumors is approximately 7.3 cases per 100,000 people [78]. Gliomas, neuroectodermal tumors derived from glial cells, encompass more than 70% of all tumors affecting the central nervous system and can be classified according to histological subtype or tumor grade (Figure 8). Most tumors arise in the subcortical region of the cerebral hemisphere. According to Furnari, “The common gliomas affecting the cerebral hemispheres of adults are termed diffuse gliomas due to their propensity to infiltrate early and extensively throughout the brain parenchyma [79].” When classified by histological appearance there are three main subtypes: oligodendrogliomas and oligoastrocytomas, or tumors which possess morphological features of both. Additionally, the World Health Organization (WHO) classifies gliomas in one of four categories based on the degree of malignancy [80]. Grade I are slow growing and benign pilocytic astrocytomas, grade II are invasive diffuse astrocytomas, and grade III are anaplastic astrocytomas. Both grades II and III possess the ability to progress to a higher grade, with grades III and IV considered to be malignant, and grade IV, glioblastoma multiforme (GBM), being the most aggressive and infiltrative (Figure 9).

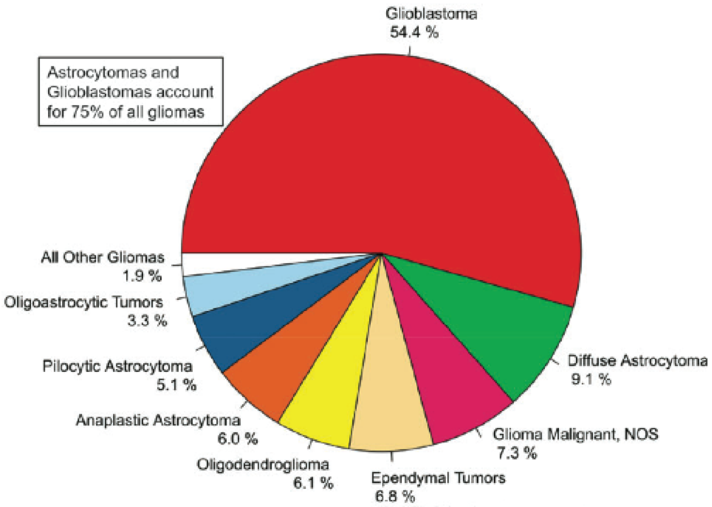


Figure 8. Distribution of primary brain and CNS cancers.
In terms of distribution by histological subtype and WHO classification, gliomas account for 75% of all brain cancers within the United States. Within this subgroup of primary brain and CNS cancers, glioblastomas account for more than 50% of all subtypes. Gliomas include: diffuse astrocytomas, oligodendrogliomas, anaplastic astrocytomas, and glioblastomas. Distribution based on study with N = 92,504 [78].

GBM can present in one of two ways: primary lesions arise de novo, whereas secondary formation involves progressive transformation from a pre-existing lower grade II diffuse astrocytoma, or grade III anaplastic astrocytoma, to a higher grade lesion. Furnari states that GBM are particularly difficult to treat given their “significant intratumoral heterogeneity on the cytopathological, transcriptional, and genomic levels [79].” These diffuse, intracranial tumors almost always recur following surgery, and are characterized by altered cellular differentiation, apoptotic resistance, necrotic lesions, genomic instability, and robust neovascularization.

Angiogenesis or neovascularization is a well-known hallmark of tumorigenesis, and many studies have demonstrated that it is crucial for solid tumor growth [81, 82]. As lesions progress from lower to higher grade, increased microvascular density, or the creation of a dense, tortuous network of microvessels, is a common feature [81, 83]. A pathological phenotype of both primary and secondary glioblastomas is microvascular hyperplasia. It consists of rapidly proliferating endothelial cells that emerge from normal microvessels as microaggregates, carrying with them components of the basal lamina necessary for angiogenesis [84]. This abnormal tumor vasculature greatly restricts oxygen delivery and also negatively impacts the efficacy of radiation therapy. Among solid tumors, glioblastomas in particular are highly vascularized, and there are high levels of the protein, vascular endothelial growth factor (VEGF), expressed in both the serum and the lesion [85]. VEGF stimulates the growth of new blood vessels and maintains tumor vasculature [86]. The overexpression of VEGF is usually indicative of a shorter survival rate [87], and it is also thought to be a contributing factor in the vascular cerebral edema observed in patients with brain cancer [88].

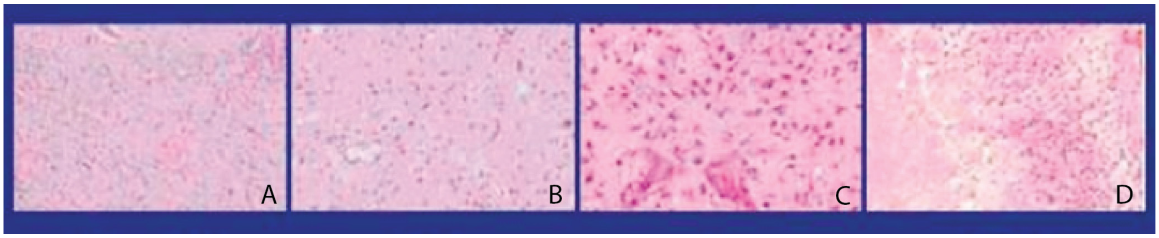


Figure 9. Glioblastoma multiforme development pathway.
The neoplastic transformation from disease-free tissue to malignant glioma is displayed: (A) normal glial tissue, (B) low grade I or grade II astrocytoma, (C) grade III anaplastic astrocytoma, (D) grade IV glioblastoma multiforme [89].

Although phenotypically similar, primary and secondary GBM associate with different genetic aberrations and arise from distinct molecular pathways, despite their indistinguishable histology. Primary GBM typically occur in older populations, with individuals above the age of forty-five, and associate with mutations in *PTEN* (phosphatase tensin homolog), a tumor suppressor gene often inactivated in cancer. Additionally, loss of heterozygosity across the majority of chromosome 10, is also observed for primary GBM, as well as amplification of *MDM2,* an E3 ubiquitin ligase whose upregulation allows for escape from p53 suppression, and *EGFR,* a receptor tyrosine kinase that controls cell proliferation [90].

Secondary GBM affect younger populations below the age of forty-five, associate with a loss of heterozygosity on the long arm of chromosomes 10, 19, and 22, or the loss of tumor suppressor genes on chromosomes 9p or 13q, as well as *TP53* mutations [90-93]. *TP53* is a tumor suppressor gene whose normal response to DNA damage involves cell cycle arrest and triggering apoptosis, but mutations render it inactive. Additionally, hypermethylation of CpG islands on cancer-specific gene promoters can result in transcriptional silencing of the associated gene. This CpG island methylator phenotype

(CIMP) is associated with glioma subtypes [94], with secondary glioblastoma displaying a higher frequency of methylation than any other glioma subtype [95-97].

Despite advances with surgical resection and radiation therapy, barriers to effective treatment still exist. Treatment options vary depending on tumor subtype and whether the disease is newly diagnosed or recurrent (Table 2). The standard of care for newly diagnosed anaplastic astrocytomas and anaplastic oligodendrogliomas, involves maximal surgical resection followed by radiation therapy, and in some instances, chemotherapy is offered as an alternative approach [98]. It is interesting to note that patients with anaplastic oligodendrogliomas and anaplastic oligoastrocytomas, that also present with codeletion of chromosome arms, 1p and 19q, respond better to treatment and are more sensitive to chemotherapy in comparison to those without this codeletion [99]. Treatment for newly diagnosed glioblastomas involves maximal surgical resection, followed by post-operative radiation, and chemotherapy with temozolomide (TMZ) [98, 100]. TMZ is an alkylating agent that adds methyl groups to DNA purine bases, thereby damaging DNA, and eventually triggering an apoptotic response. Patients treated with TMZ following radiation therapy have prolonged survival rates and delayed tumor progression, in comparison to those only treated with radiation [98, 100].

Complete surgical resection is not always possible given the infiltrative nature of gliomas and there is a high rate of recurrence among patients. A PCV combination regimen can be administered in cases of recurrent anaplastic oligodendroglioma. However, the chemotherapeutic agent most commonly used in cases of recurrent disease is TMZ. It can be dispensed to patients with recurrent anaplastic astrocytoma, anaplastic oligodendroglioma, or GBM, as an injection or oral therapeutic.

Table 2. Summary of current treatment options for malignant gliomas.

Tumor Type	**Standard Treatment**	**Alternatives**
Newly diagnosed glioma		
Anaplastic oligodendroglioma	Maximal surgical resection; RT	TMZ or PCV
Anaplastic astrocytoma	Maximal surgical resection; RT	RT/TMZ; Carmustine/RT
Glioblastoma	Maximal surgical resection; RT; TMZ	Modified RT; or RT/TMZ
Recurrent disease		
Anaplastic oligodendroglioma	Reoperation in selected patients or PCV; concurrent TMZ	Irinotecan; Erlotinib; Gefitinib; or Imatinib
Anaplastic astrocytoma	Reoperation or TMZ	PCV
Glioblastoma	Reoperation or TMZ	Procarbazine

Abbreviations: RT, radiation therapy; PCV regimen, procarbazine, lomustine, vincristine.

The survival rates for individuals with malignant gliomas, especially those with GBM, remain particularly low and are associated with disproportionately high morbidity and mortality (Figure 10). The median survival rate for patients with glioblastomas is 9-15 months with less than 5% surviving for more than five years [98, 101]. This survival rate is only slightly longer for those with anaplastic gliomas, 2-5 years [100, 101]. Concurrent radiation therapy and TMZ treatment has had a certain measure of success; however, even treatment therapy with TMZ is subject to tumor resistance [102, 103]. The

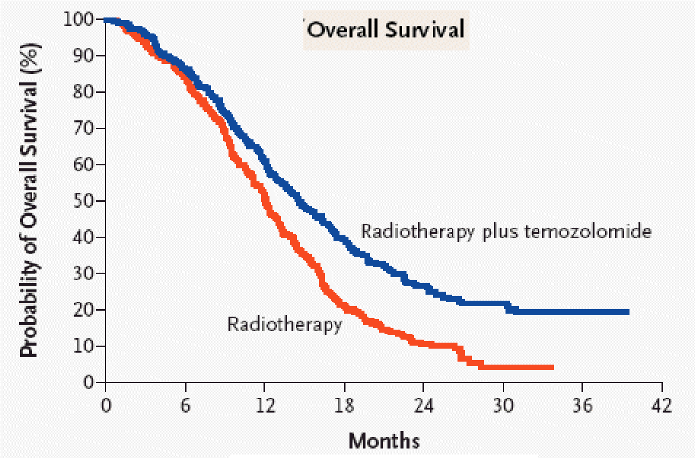


Figure 10. Glioblastomas are associated with high morbidity and mortality. Overall survival in trial of temozolomide in conjunction with radiation therapy (TMZ/RT) versus RT alone. Randomized phase III trial comparing initial therapy of RT alone with RT and concomitant maintenance with TMZ [100].

methylguanine methyltransferase (MGMT) is a DNA excision repair enzyme that targets and removes alkyl groups that have been added to the guanine bases of DNA. When MGMT is amplified and overexpressed in gliomas it can reverse the effects of TMZ [103]. Although progress has been made with tumor detection, surgical techniques, radiosensitizers, angiogenesis inhibitors, chemotherapeutic drugs, and gene therapy, no single class of agents possessing superior efficacy against gliomas has emerged. Of late, studies are focusing more attention on finding new targets for treatment, by obtaining a better understanding of the molecular pathogenesis of this disease.

2.4 Role of isocitrate dehydrogenase mutations in cancer

An initial study in 2006 identified somatic mutations of *IDH1* in colorectal cancer [104]. A subsequent study by Parsons *et al.* uncovered somatic mutations in *IDH1* occurring in over 12% of glioblastomas during a genome-wide mutational analysis screen. Sequencing analysis of exon 4 revealed a recurrent point mutation at codon 395 in which adenine was substituted for guanine (G395A), resulting in the replacement of arginine with a histidine at amino acid 132 (R132H), a residue thought to be important for catalysis. This was the most prominent amino acid substitution with other substitutions (R132C, R132S) occurring less frequently [105]. Additional sequencing efforts have uncovered aberrations in *IDH* that resulted in other malignancies (Table 3), and these include: T-cell lymphomas, prostate carcinomas, cholangiocarcinomas, pheochromocytomas, enchondromas, paragangliomas, spindle cell hemagiomas, chondrosarcomas, B-cell acute lymphoblastic leukemias, colon cancer, myeloproliferative neoplasms, melanomas, and thyroid carcinomas [106-112].

2.4.1 Prevalence of *IDH* mutations in gliomas and acute myeloid leukemia

Mutations in *IDH1* or *IDH2* are highly associated with gliomas and acute myeloid leukemia (AML). All mutations share these common features: somatic, monoallelic with one wild type allele being retained, and involve the substitution of an arginine residue at a recurrent position. Studies with large cohorts underscored a high prevalence of *IDH* mutations associating with gliomas and AML. They occur in 50% - 94% of grade II and grade III gliomas, and in 73% - 88% of secondary glioblastomas [107, 113-119]. Mutant IDH1 has also been identified to a lesser extent in 8% of acute myeloid leukemia (AML) cases [118].

In other cancer genome studies, *IDH2* mutations were found to be more common in AML than gliomas, occurring in 20% of cases of AML [120]. Individuals lacking point mutations in *IDH1* at R132 will often have analogous mutations in *IDH2* at R172, with R172K being the most prominent substitution and R172W occurring far less frequently. There is a hot spot on *IDH1* at R132 and another hot spot (Figure 11) on *IDH2* at R172

[110, 115, 120-124]. *IDH2* mutations at R140 have also been reported with R140Q being the most prominent (95%), while R140L and R140W substitutions are less common [120]. Mutations at this position only occur in this tumor type.

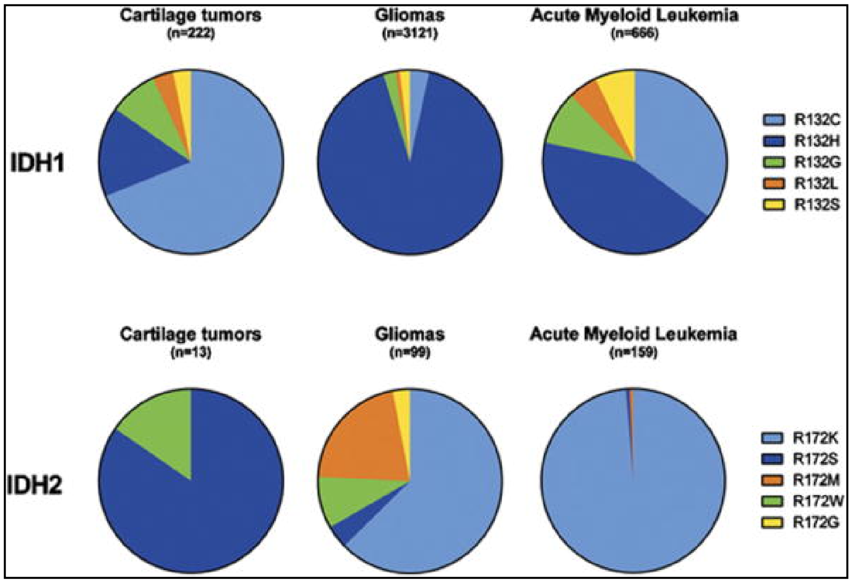


Figure 11. Mutations in *IDH1* and *IDH2* are prevalent in gliomas and AML. Mutations of *IDH1* are more prevalent in gliomas while mutations in *IDH2* are more prominent in cases of AML. There is a hot spot located at R132H on *IDH1* and an analogous hot spot at R172K located on *IDH2*. [125].

IDH mutations may predispose cells to become malignant since they occur before the development of other aberrations [117, 126]. Histological analysis of paired samples in glioma patients transforming from low-grade to high-grade lesions showed aberrations in *IDH* were early events in tumor progression [118]. Surprisingly, mutations in *IDH1* or *IDH2* have not been observed in primary GBM, and are generally not seen in other cancers of the central nervous system [114, 116, 117]. *IDH* mutations often segregate with loss of chromosomal arms 1p and 19q in the majority of diffuse gliomas [127]. As it relates to clinical outcomes, individuals with glioblastomas harboring mutant forms of

IDH tend to have a better prognosis than those with wild type IDH (Figure 12). Overall, their median survival rate tends to be longer [116, 119, 128-130]. Although these prognostic results are interesting, there are no reports that clearly define *IDH* mutations as significant predictive indicators of disease progression or lack thereof.

2.4.2 Neomorphic activity of Human IDH mutants

Mutant forms of IDH1 and IDH2 lose their ability to convert isocitrate to αKG through oxidative decarboxylation. Following this observation, Yan and colleagues were the first to assert that the R132H mutation in *IDH1* and the R172K mutation in *IDH2* result in a loss of enzymatic function [119]. Subsequent studies verified this loss of function, and suggested that IDH1 was a tumor suppressor gene displaying dominant negative activity, as mutant IDH1 formed heterodimers with wild type IDH1. [131].

Zhao postulated that the formation of catalytically inactive heterodimers could lead to the inactivation of IDH1, causing a reduction in αKG levels within the cell, and inducing HIF1-α to activate target genes and drive tumorigenesis. However, this hypothesis was challenged by other reports that demonstrated that mutant heterodimers (WT:R132H) do not exhibit dominant negative inhibition, nor are they formed in excess of wild type or mutant homodimers [132, 133]. Additionally, these mutations were not observed to be sufficient to upregulate HIF1-α. Dang and colleagues put forward another hypothesis and reported a novel gain of function [43].

IDH mutants possess neomorphic enzyme activity (Figure 13) and catalyze the reduction of αKG to 2-hydroxyglutarate (2HG) [43]. Given this gain of function, *IDH1*

Table 3. Somatic *IDH1* and *IDH2* mutations were identified in different tumors.

Tumor Type	***IDH1***	***IDH2***	**References**
Glioma	R132H	R172K	[43, 105, 113, 114, 116, 117, 119]
	R132C	R172M	
	R132S	R172S	
	R132G	R172G	
	R132L	R172W	
	R132V		
Acute myeloid leukemia	R132H	R140Q	[110, 115, 118, 120-124]
	R132G	R172K	
	R132C	R140W	
	R132S	R172G	
	R132L	R140L	
	V71I		
Prostate carcinoma	R132H		[107]
	R132C		
B-cell acute lymphoblastic leukemia	R132H		[107]
Paraganglioma	R132C		[112]
	R132H		
Colon cancer	R132C		[104, 107, 114]
Thyroid carcinoma	V71I		[111]
	G123R		
Chondrosarcoma	R132C	R172S	[106]
	R132H		
	R132G		
Enchondroma	R132H	R172K	[109]
Cholangiocarcinoma	R132C	R172W	[108, 134]
	R132L		
	R132G		

behaves more like an oncogene than a tumor suppressor. This study was the first to report the production of the oncometabolite, 2HG, accumulating to unusually high, millimolar concentrations within cells, from the NADPH-dependent reduction of αKG [43]. This and subsequent studies demonstrated that IDH1 mutants (R132H, R132C, R132L, R132S, R132G, R132V) and IDH2 mutants (R172K, R172M, R172S, R172G, R172W) all possess the ability to reduce αKG to 2HG [43, 121, 122, 124, 132].

Different research groups performed mass spectrometry analysis and were able to confirm 2HG as the direct product of these mutant enzymes [43, 122, 124]. Vogelstein and colleagues regard IDH as an oncogene and more specifically, a Mut-driver gene. Mut-driver genes contain mutations that directly or indirectly confer a selective growth advantage to cells [60]. Possible mechanisms are emerging that connect the accumulation of 2HG with impaired dioxygenase activity and altered epigenetic regulation.

2.4.3 Tumorigenic potential of 2-hydroxyglutarate accumulation

Although countless cancer genome studies identified mutations in *IDH* genes, only recently has the underlying mechanism of how these mutations drive tumorigenesis been described. The metabolic product of mutant IDH is 2-hydroxyglutarate (2HG) and structurally, 2HG and αKG are nearly identical. The only exception is that the carbonyl group at the C2 position in α-ketoglutarate is replaced by a hydroxyl group in 2HG. This oncometabolite competes with αKG-dependent dioxygenases for cytosolic stores of αKG, and this superfamily of dioxygenases is responsible for catalyzing several different reactions within the cell [29]. Mutations in IDH1/IDH2 promote tumorigenesis through the inhibition of dioxygenases such as TET2 and KDM [135].

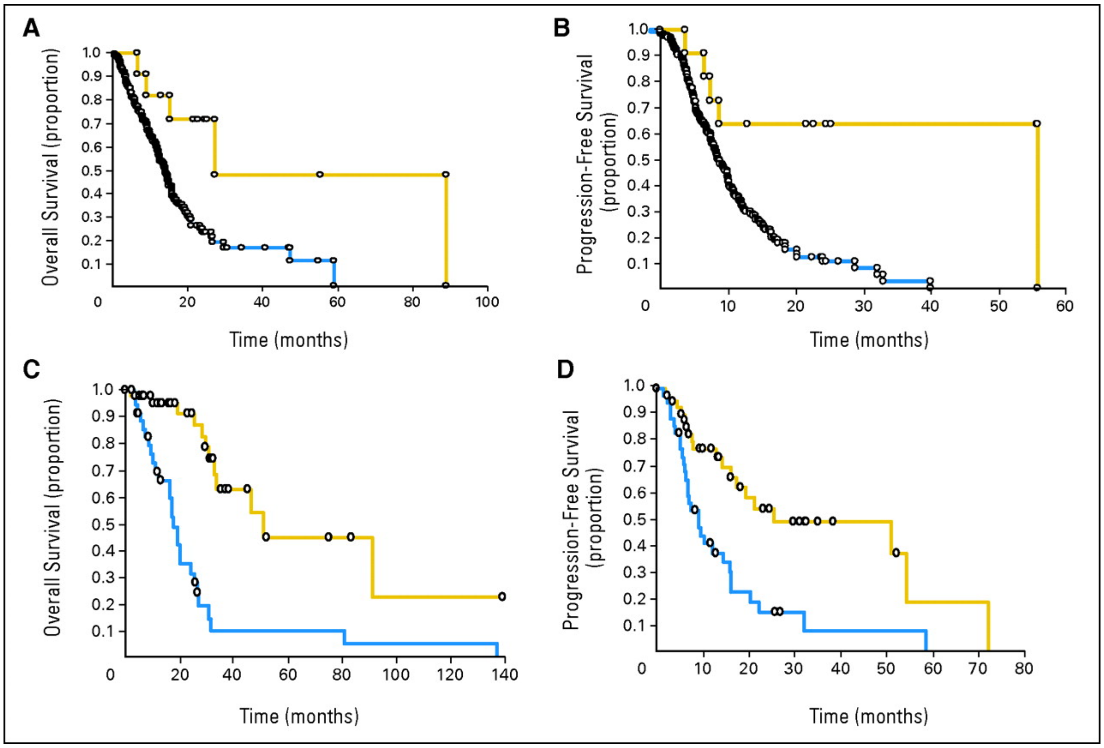


Figure 12. Patients with IDH1 mutation have better prognosis.
(A) Overall survival curves and (B) progression-free survival curves in glioblastoma according to IDH1 codon 132 status (gold line, mutated; blue line, nonmutated). (C) Overall survival curves and (D) progression-free survival curves in grade 3 gliomas without 1p19q codeletion and no EGFR amplification according to IDH1 codon 132 status (gold line, mutated; blue line, nonmutated) [136].

2HG exerts tumorigenic potential (Figure 14) by impairing the activity of prolyl hydroxylases, the TET family of DNA hydroxylases, and members of the Jumonji family of histone lysine demethylases (KDM). It has deleterious effects on several processes within the cell, including the hypermethylation of DNA and chromatin [135, 137]. These alterations cause epigenetic reprogramming, and ultimately lead to tumorigenesis. Additionally, prolyl hydroxylases (PHDs) regulate the activity of HIF proteins by adding hydroxyl groups to key proline residues. The stability and activity of HIF1α is also affected by ROS and intracellular levels of αKG [138]. There is some debate as to whether abnormal accumulation of 2HG can also cause the inhibition of HIF prolyl hydroxylases and lead to the stabilization of HIF1α. Studies probing the effects of 2HG accumulation on PHDs and HIF activity have yielded conflicting *in vitro* results. Some studies demonstrate that 2HG acts as a PHD inhibitor and results in the expression of more HIF protein, with less targeting of HIF protein for proteasomal degradation [137]. Still other reports provide evidence that 2HG acts as a PHD activator, and increases the levels of HIF protein [139, 140].

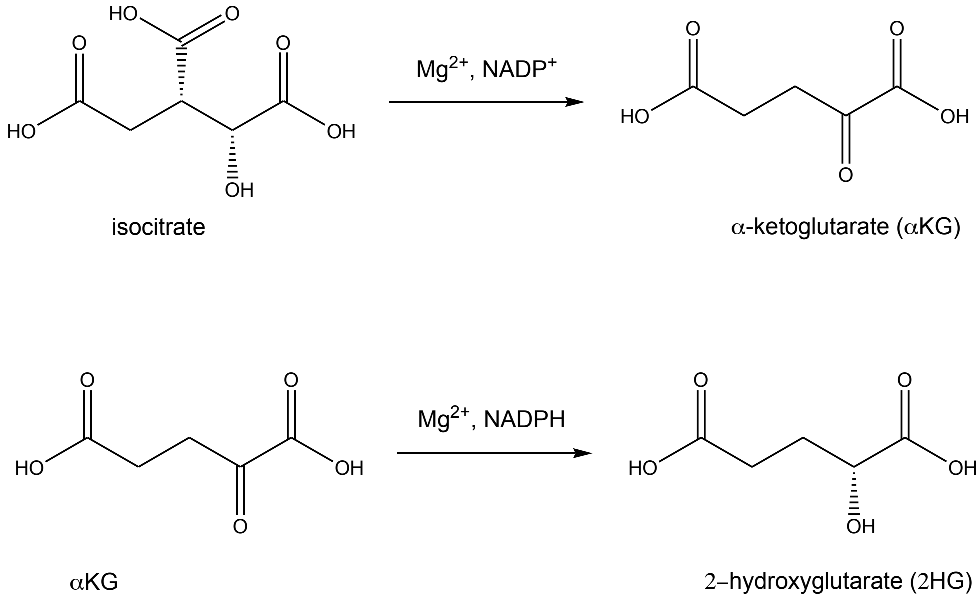


Figure 13. Reaction catalyzed by mutant IDH enzymes.
(A) Reaction catalyzed by wild-type IDH. (B) Reaction catalyzed by mutant IDH.

TET enzymes catalyze the hydroxylation of 5-methylcytosine residues in DNA. The hydroxylation of these cytosine residues ultimately results in global DNA remodeling. When TET2 is inhibited, it prevents this intermediate step from occurring, and affects gene transcription and genome stability. Data also suggests that 2HG accumulation inhibits the activity of histone demethylases such as KDM, enzymes that regulate patterns of histone methylation at lysine residues, and causes chromatin remodeling [141]. Overall, the inhibition of these dioxygenases results in alterations that block the cellular differentiation of neuronal cells and hematopoietic cells, cause epigenetic reprogramming, and ultimately lead to tumorigenesis [140-142].

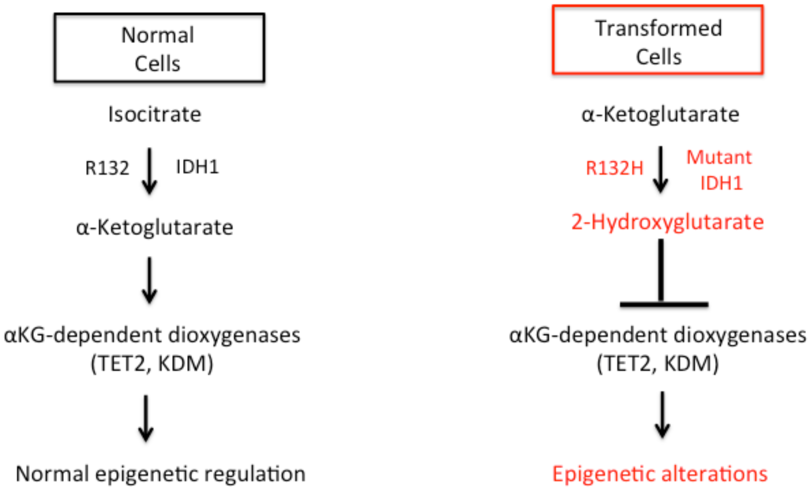


Figure 14. The oncometabolite 2HG inhibits αKG-dependent dioxygenases. IDH mutants produce 2HG from NADPH-dependent reduction of αKG. Tumor cells expressing mutant IDH1 or IDH2 have high levels of 2HG. 2HG is an oncometabolite that inhibits αKG-dependent dioxygenases, alters DNA and histone methylation, and causes to epigenetic alterations that ultimately lead to tumorigenesis.

2.5 Diagnostic and therapeutic targeting of Human IDH mutation

Quite often *IDH* mutations converge with other genetic aberrations to contribute to the malignant transformation of brain tissue. Gliomas possess different molecular signatures: 1p/19q codeletion, MGMT promoter methylation, *IDH* mutation, and glioma-specific CIMP, as well as mutations in *PTEN*, *MDM2*, *EGFR*, and *TP53*. The impact of these molecular alterations on patient outcomes and therapy is still being assessed. However, as treatment options evolve, and we learn more about the molecular underpinnings of gliomas and their tumor metabolism, the possibility exists that the use of molecular biomarkers such as IDH1, in addition to histological grade and subtype, will be included in the development of new treatment regimens based on the molecular status of the tumor.

The detection of *IDH* mutations through the use of a monoclonal antibody that recognizes IDH1 R132H protein, is already being utilized as a diagnostic marker in the assessment of IDH status in glioma biopsy sections [143]. In the future, the possibility also exists that 2HG might be used as a biomarker by clinicians. It can be detected in the serum of patients and could serve as a less invasive means to track disease progression or recurrence than a biopsy. It could also serve as a means of tracking the effectiveness of a new treatment regimen.

Small molecules that specifically target the enzymatic activity of mutant IDH1 and IDH2 would also provide a new therapeutic opportunity for drug development. X-ray crystallography is a powerful tool that provides detailed three-dimensional information about binding pockets. In the past, structure-based drug design (SBDD) has been used with some success to find small molecules when the structure of a target is known [144, 145]. Bristol-Myers Squibb developed Captopril, and it was among the first successes of SBDD in the late 70s. The development of HIV inhibitors followed in the early 90s. Knowledge of the ATP-binding site of kinases led to the discovery of the cancer therapeutics Erlotinib (Tarceva) and Imatinib (Gleevac). Modeling and virtual screening of binding sites can assist with drug design by mapping the spatial and physico-chemical properties of the binding pocket. More recently, fragment-based drug design is emerging as a new screening tool that can lead to the rational design and synthesis of lead compounds [146].

2.5.1 Protein-Protein Interfaces and Structure-Based Drug Design

Protein-protein interfaces (PPI) have not been the first choice for drug developers, mostly due to the difficulty in targeting them. The incentive for drug discovery teams to push the boundary comes from the saturation of established drug discovery targets such as protein kinases, nuclear receptors, voltage-gated ion channels, and G protein-coupled receptors. PPIs reveal a new and virtually inexhaustible area for drug discovery. The success of several drugs that target PPIs demonstrates that they are the next frontier [147, 148]. In the last few years the inhibitors Navitoclax and Obatoclax have been found to target interactions of BCL-1 with the tumor suppressors p53 and MDM2 [149]. Additionally, a drug for dry eyes that targets the interaction between lymphocyte function-associated antigen-1 and the intercellular adhesion molecule-1 (LFA-1/ICAM) has entered Phase III clinical trials. The determination of binding sites that contribute to the free energy of interaction has also encouraged the development of tethering methods that target such sites. In most cases, structural biology is used to determine the site of the PPI, but in some cases mutagenesis experiments also complement and assist in localizing the amino acid residues that contribute to the binding energy.

2.5.2 Classes of Protein-Protein Interactions

There are two types of protein-protein interactions that result in the formation of either obligate, stable complexes, or transient complexes [150]. The latter are the consequence of a signaling or activation process, or posttranslational modification and involve binding cofactors and substrates. The interface formed by obligate complexes is large and hydrophobic, while transient complexes tend to have a more polar surface area [150]. Within this vast landscape of overlapping interfaces, there are interactions that occur between protein complexes that can be targeted and trapped during these conformational changes. Quite often, these interactions occur at a dimer interface as in the case of survivin [151, 152]. The binding of cAMP-response-element-binding protein (CREB) to CREB-binding protein (CBP) is an example of folding and binding where the groove of the kinase inducible interacting domain of CBP accommodates the helices of the kinase inducible domain of CREB [153]. The NMR structure gave information about the binding interactions between CBP and CREB, and this interface has been disrupted by the compound KG-501 [154].

2.5.3 Physical characteristics of new PPI sites

Protein-protein interfaces are often characterized by a large surface area of 1500-3000 Å^2 [155, 156]. Most effective small molecules target a smaller area rather than the whole interface. The molecular surface tends to be flatter than a typical active site, and has 'energetic hot spots' [157]. However, most surfaces have an overrepresentation of hydrophobic residues clustered in patches. Van der Waal's forces, hydrogen bonding and charge-charge interactions are the obvious weak electronic interactions observed. The compounds that disturb PPI tend to be large and lipophilic, with several aromatic rings and fewer rotatable bonds. In addition, these compounds tend to be more rigid than those that typically bind active sites.

2.5.4 Discovering and navigating new PPI landscapes

Small molecule and fragment-based screening can be successfully used to explore the chemical space efficiently [158, 159]. For example, chemical libraries of low affinity that can be linked or tethered to protein targets are employed. The use of mass spectrometry to detect compounds with an active thiol is a method engineered by Wells [160]. The limitation of this method is the requirement of engineered cysteines in the protein target. A second approach is to dock compounds by conventional fragment based screening using libraries [159, 161]. X-ray crystallography can detect even low, millimolar binders. Nuclear magnetic resonance (NMR), X-ray crystallography, surface plasmon resonance (SPR), thermal shift fluorimetry (TSF), differential scanning fluorimetry (DSF), or isothermal calorimetry (ITC) are required to screen and validate fragment binding [161]. Although *in silico* design has its limitations, it can yield valuable information about the chemical space. Structure-activity relationships are a source of information that can complement data from TSF, DSF, and ITC. Another more creative approach, involves the use of porphyrins or proteomimetics such as stapled peptides, or helical elements to probe binding interactions.

2.6 Objectives

The overall goal of this project was to advance the knowledge of the molecular architecture and reaction mechanism of the IDH1 R132H mutant through crystal structure determination. We were interested in finding small molecules that inhibit mutant IDH1 to suggest possible modifications or medicinal chemistry for the development of cancer therapeutics. Our specific aims relating to this project were twofold: we sought to explore the mechanisms of action of small molecules that interact with the IDH1 R132H mutant; we also used structural biology to observe the binding site of small molecules to the IDH1 R132H mutant to map the spatial and biophysical properties, to determine if this protein-protein interface was druggable and amenable to structure-based drug design.

3. Methods

3.1 General Methods

3.1.1 High-throughput screen of small molecules

A High-throughput screen was conducted on a 384 well plate using a subset of the ChemBridge Library. Small molecules were selected based on a three-dimensional pharmacophore in order to identify α-ketoglutarate mimetics. Forty-eight compounds with one or more carboxylic acid groups or derivatives thereof were selected and screened for inhibition of IDH1 R132H (Figure 14). From this primary screen, three compounds were selected and advanced to a biochemical activity assay. Compound concentration was set at 0.133 mM and the assay carried out in the reducing direction with 33 mM Tris acetate pH 7.4, 1.3 mM $MgCl_2$, 40 mM $NaHCO_3$, 0.4 mM NADPH, 0.5 mM α-KG and purified IDH1 R132H (Figure 15). Assay progression was monitored at 340 nm using a microplate reader (BMG POLARstar Omega).

3.1.2 Enzymatic inhibition assays with small molecules

The consumption of NADPH was used to measure the activity of IDH1 R132H in the presence of small molecules, as it converted α-ketoglutarate to 2-hydroxyglutarate. Small molecules, A07 (ChemBridge, 5662524) and ZAG (XcessBiosciences, IDH-C35), (Figure 16) were prepared as 10 mM stocks using DMSO and subsequently diluted to a concentration of 0.2 mM in the final reaction volume. The assay was conducted in a 384 well plate with 50 μL reactions. Reactions were carried out with 100 μg/mL of IDH1

R132H diluted in assay buffer containing 20 mM Tris-Cl pH 7.5, 150 mM NaCl, 10 mM $MgCl_2$, 0.200 mM NADPH, 0.05% bovine serum albumin, and 2 mM β-mercaptoethanol in a volume of 39 μL, with the addition of 1 μL of 10 mM stock of A07 or 10 mM stock ZAG, and incubated for 60 minutes. Additionally, 40 μL control reactions containing assay buffer with and buffer without NADPH (20 mM Tris-Cl pH 7.5, 150 mM NaCl, 10 mM $MgCl_2$, 0.05% bovine serum albumin, and 2 mM β-mercaptoethanol) were prepared as outlined above. These reactions with IDH1 R132H alone, A07 alone, and ZAG alone were utilized to analyze the absorbance of the compound both in the presence and absence of NADPH. Following the addition of 10 μL of 5 mM αKG (the substrate), IDH1 R132H activity was monitored spectrophotometrically at 340 nm using a microplate reader (BMG POLARstar Omega). In order to assess the consumption of IDH1 R132H over time, scans were taken before the addition of substrate, and after at 2, 15, 30, 60, 90, 120, 180, and 240 minutes, in the absence or presence of 0.2 mM A07 or 0.2 mM ZAG.

3.1.3 Sequence alignment of IDH enzymes

The amino acid sequences of three eukaryotic species of isocitrate dehydrogenase and two human isoforms of IDH were selected for structure-guided sequence alignment. The selected protein sequences were aligned using Clustal Omega [162]. The resulting figures containing secondary structural elements were generated using the program ESPript [163].

3.2 Methods for structure determination of selenomethionine-replaced IDH1 R132H

3.2.1 Protein expression and purification

Full-length *Homo sapiens IDH1* was cloned into the pET-22a vector and mutant *IDH1* R132H was generated using QuickChange site-directed mutagenesis (Stratagene). Both plasmids were a generous gift from the laboratory of Dr. Bert Vogelstein (Ludwig Center for Cancer Genetics and Therapeutics). Plasmid DNA was used to transform bacterial cells and IDH1 R132H was expressed in *E. coli* B834 (DE3) strain (Novagen) to generate the seleno-methionine replaced protein. Cells were grown overnight at 37°C in 60 mL M9 media supplemented with 5% LB media and 100 μg/mL ampicillin. Cells were harvested at 6000 rpm in a Sorvall GSA rotor at 4°C, and resuspended in M9 media containing 100 μg/mL seleno-L-methionine (Sigma) and 100 μg/mL ampicillin. After growth to 0.6 at A_{600}, expression was induced with 1 mM β-D-1-thiogalactopyranoside (IPTG) and growth continued overnight at 25°C. Cells were harvested by centrifugation, washed with 20 mM Tris-Cl pH 7.4, 500 mM NaCl, 10 mM imidazole, 10% glycerol, 0.1% Triton X-100, 5 mM β-mercaptoethanol, pelleted, and stored at −80°C.

After thawing, cells were resuspended in lysis buffer: 20 mM Tris-Cl pH 7.4, 500 mM NaCl, 10 mM imidazole, 10% glycerol, 0.1% Triton X-100, and 5 mM β-mercaptoethanol, with one dissolved Complete EDTA-free protease inhibitor tablet (Roche), and lysed using microfluidization. Cellular debris was pelleted and removed by centrifugation on a Sorvall GSA rotor for 30 minutes at 12,000 rpm. The lysate was filtered through a 0.2 μm filter and loaded onto a HisTrap HP 5 mL column (GE Healthcare) pre-equilibrated with buffer containing 20 mM Tris-Cl pH 7.4, 500 mM NaCl, 10 mM imidazole, 10% glycerol, and 5 mM β-mercaptoethanol. Unbound proteins were washed for 20 column volumes with this buffer. Target protein was eluted by an imidazole gradient with 10-50% of buffer containing 20 mM Tris-Cl pH 7.4, 500 mM NaCl, 500 mM imidazole, 10% glycerol, 5 mM β-mercaptoethanol (Figure 20). Fractions containing protein were pooled and dialyzed using 10 kD MWCO membrane (Pierce Snakeskin) against an excess of buffer containing 20 mM Tris-Cl pH 8.5, 150 mM NaCl, and 5 mM β-mercaptoethanol, at 4°C with agitation overnight. The following morning, protein was dialyzed twice against 2L of binding buffer (20 mM Tris-Cl pH 8.5, 50 mM NaCl, and 5 mM β-mercaptoethanol) for a total of 4 hours. The dialysis buffer was filtered through a 0.2 μm filter and used as the equilibration and binding buffer for the next step of purification. The protein was further purified over a Source Q column (GE Healthcare), using anion exchange chromatography. The column was pre-equilibrated with buffer A (20 mM Tris-Cl pH 8.5, 50 mM NaCl, and 5 mM β-mercaptoethanol) before applying protein sample over the column. Unbound proteins were washed with 4 column volumes of binding buffer, and target protein was eluted with a gradient to 50% of buffer B (20 mM Tris-Cl pH 8.5, 1 M NaCl, and 5 mM β-mercaptoethanol) over 20 column volumes (Figure 21). Fractions of protein containing ≥ 95% purity were collected and dialyzed against an excess of buffer containing 20 mM Tris-Cl pH 8.0, 150 mM NaCl, and 5 mM β-mercaptoethanol. Target protein was concentrated to 15-20 mg/mL, and stored at −80 °C. All chromatographic separations were performed on AKTA Basic Fast Protein Liquid Chromatography (FPLC) system.

3.2.2 Crystallization of selenomethionine-replaced IDH1 R132H

Crystallization screening of Se-Met IDH1 R132H was performed using incomplete factorial screens. Selected conditions were optimized using hanging drop vapor diffusion with a reservoir of 500 μL. Protein complexes were co-crystallized from solutions consisting of 15-20 mg/mL Se-Met IDH1 R132H mixed with either 1 mM αKG or 1 mM ICT. Drops containing 1 μL of protein and 1 μL of reservoir solution were equilibrated against a reservoir containing 0.1 M Tris-Cl pH 8.0, 0.2 M ammonium sulfate, and 25.0 - 27.5% PEG 3350 (w/v). Diffraction quality crystals were obtained in 2 weeks at 20°C. Crystals were soaked with 0.02 mM A07 for 60 min and frozen using 2 M lithium sulfate as a cryoprotectant.

3.2.3 MAD data collection and processing

The X-ray diffraction data for selenomethionine-derivatized IDH1 R132H crystals and selenomethionine-derivatized IDH1 R132H crystals in complex with inhibitor A07 were collected at beam line X6A of the National Synchotron Light Source (NSLS) at Brookhaven National Laboratory, at the selenium peak wavelength of 0.9793 Å. The crystals diffracted to a resolution of 2.1 Å in the absence of inhibitor and 3.0 Å in the presence of A07 respectively (Table 4). Data were processed with HKL2000 [164] .

3.2.4 Structure determination, phase calculation, and refinement

Phases were determined by multiple wavelength anomalous diffraction (MAD) using data collected on selenomethionine-derivatized crystals of IDH1 R132H. The SOLVE/RESOLVE software suite was used to determine the positions of the anomalously diffracting selenium atoms and for initial model building [165-168]. The structure of the selenomethionine-derivatized IDH1 R132H crystals, in complex with ICT-NADP and A07, were determined by Fourier synthesis using the coordinates of the selenomethionine IDH1 model. After rigid body and positional refinement, iterative rounds of model rebuilding were carried out with the program Coot [169]. The initial model was refined using PHENIX [170, 171] or REFMAC5 [172] in the CCP4i suite. A free R factor set of 5 % of randomly chosen reflections was kept during the refinement for cross validation purposes. Figures of the structures were prepared using the program Pymol [173]. Refinement statistics for both data sets are given in Table 4.

3.3 Methods for structure determination of IDH1 R132H

3.3.1 Protein expression and purification

Plasmids containing full-length *H. sapiens IDH1* and *IDH1* R132H were a gift from the laboratory of Dr. Bert Vogelstien. Plasmid DNA of *IDH1* R132H was used to transform BL21 (DE3) cells (Novagen). Eight liters of bacterial cells were grown at 37°C in LB medium supplemented with 100 μg/mL ampicillin to 0.7 at A_{600}. Protein expression was induced with the addition of 1 mM IPTG and growth was continued at 18°C overnight. Cells were harvested by centrifugation and washed with 20 mM Tris-Cl pH 7.4, 500 mM NaCl, 10 mM imidazole, 10% glycerol, 0.1% Triton X-100, and 5 mM β-mercaptoethanol. Following this wash step, cells were subsequently pelleted and stored at −80 °C.

Cells were resuspended in lysis buffer containing 20 mM Tris-Cl pH 7.4, 500 mM NaCl, 10 mM imidazole, 10% glycerol, 0.1% Triton X-100, 5 mM β-mercaptoethanol, with one dissolved Complete EDTA-free protease inhibitor tablet (Roche) after thawing. Following lysis via microfluidization, cellular debris was pelleted and removed by centrifugation in a GSA rotor for 30 minutes at 12,000 rpm. The lysate was filtered through a 0.2 μm filter and loaded onto a HisTrap HP 5 mL column (GE Healthcare) pre-equilibrated with buffer A (20 mM Tris-Cl pH 7.4, 500 mM NaCl, 10 mM imidazole, 10% glycerol, and 5 mM β-mercaptoethanol). Unbound proteins were washed for 20 column volumes with this buffer. An imidazole gradient was used to elute target protein with 10-50% of buffer B (20 mM Tris-Cl pH 7.4, 500 mM NaCl, 500 mM imidazole, 10% glycerol, and 5 mM β-mercaptoethanol) (Figure 22). Fractions containing protein were pooled and dialyzed using 10 kD MWCO membrane (Pierce Snakeskin) against an excess of buffer containing 20 mM Tris-Cl pH 8.5, 150 mM NaCl, and 5 mM β-mercaptoethanol, at 4°C with agitation overnight. The following morning, protein was dialyzed twice against 2L of binding buffer (20 mM Tris-Cl pH 8.5, 50 mM NaCl, and 5 mM β-mercaptoethanol), for a total of 5 hours. The dialysis buffer was filtered through a 0.2 μm filter and used as the equilibration and binding buffer for the next step of purification. The protein was further purified over a Source Q column (GE Healthcare), using anion exchange chromatography. The column was pre-equilibrated with buffer A (20 mM Tris-Cl pH 8.5, 50 mM NaCl, and 5 mM β-mercaptoethanol) before applying protein sample over the column. Unbound proteins were washed with 4 column volumes of binding buffer and target protein was eluted with a gradient to 100% of buffer B (20 mM Tris-Cl pH 8.5, 1 M NaCl, and 5 mM β-mercaptoethanol) over 20 column volumes (Figure 23). Fractions of target protein containing ≥ 95% purity were collected and dialyzed against an excess of buffer containing 20 mM Tris-Cl pH 8.0, 150 mM NaCl, and 5 mM β-mercaptoethanol. Protein was concentrated to 15-20 mg/mL, and stored at −80 °C.

3.3.2 Crystallization of IDH1 R132H

The crystallization screening of IDH1 R132H was performed as described for IDH1 R132H selenomethionine-derivatized protein. In short, selected conditions were optimized using hanging drop vapor diffusion, with diffraction quality crystals obtained in 2 weeks at 20°C with drops containing 1 μL of protein and 1 μL of reservoir solution, equilibrated against a reservoir of 100 mM Tris-Cl pH 8.0, 200 mM ammonium sulfate and 25.0 - 27.5% PEG 3350 (w/v). Protein complexes were co-crystallized from solutions consisting of 15-20 mg/mL IDH1 R132H mixed with 1 mM ICT (Figure 24). Crystals were soaked with 0.02 mM ZAG for 60 min and frozen using 2 M lithium sulfate as a cryoprotectant.

3.3.3 Data collection and processing

Data for IDH1 R132H crystals in complex with ZAG were collected at LRL-CAT beam line 31A of the Advanced Photon Source (APS). The crystals diffracted to a resolution of 3.3 Å in the presence of ZAG. Indexing and data reduction for crystals was carried out using HKL2000 [164].

3.3.4 Structure determination, phase calculation, and refinement

The structure of IDH1 R132H crystals in complex with ICT-NADP and ZAG, were determined by Fourier synthesis using the coordinates of the selenomethionine-derivatized IDH1 model. To summarize, Coot was used for iterative rounds of rebuilding [169], with further refinement performed with REFMAC5 [172]. A free R factor set of 5% of randomly chosen reflections was kept during the refinement for cross validation purposes. Pymol was used to generate figures of the structures [173] Refinement statistics for this data set are given in Table 4.

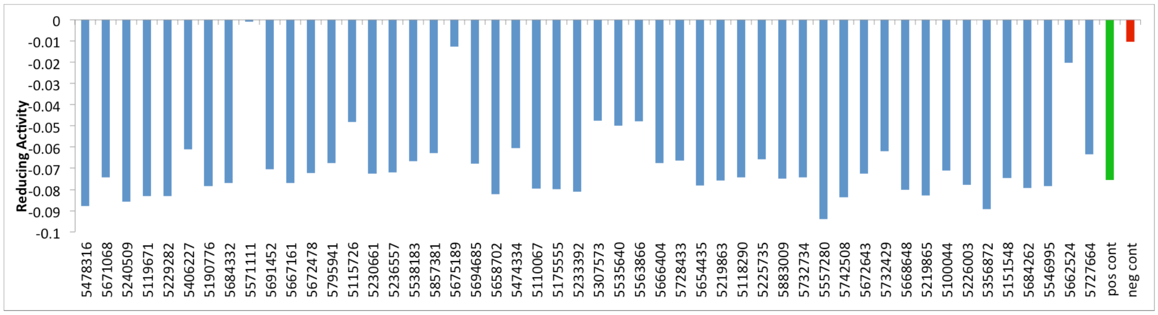


Figure 15. High-throughput screen based on three-dimensional pharmacophore. An initial screen was conducted with a subset of small molecules from the ChemBridge Library. Three small molecules were selected from this initial screen and advanced.

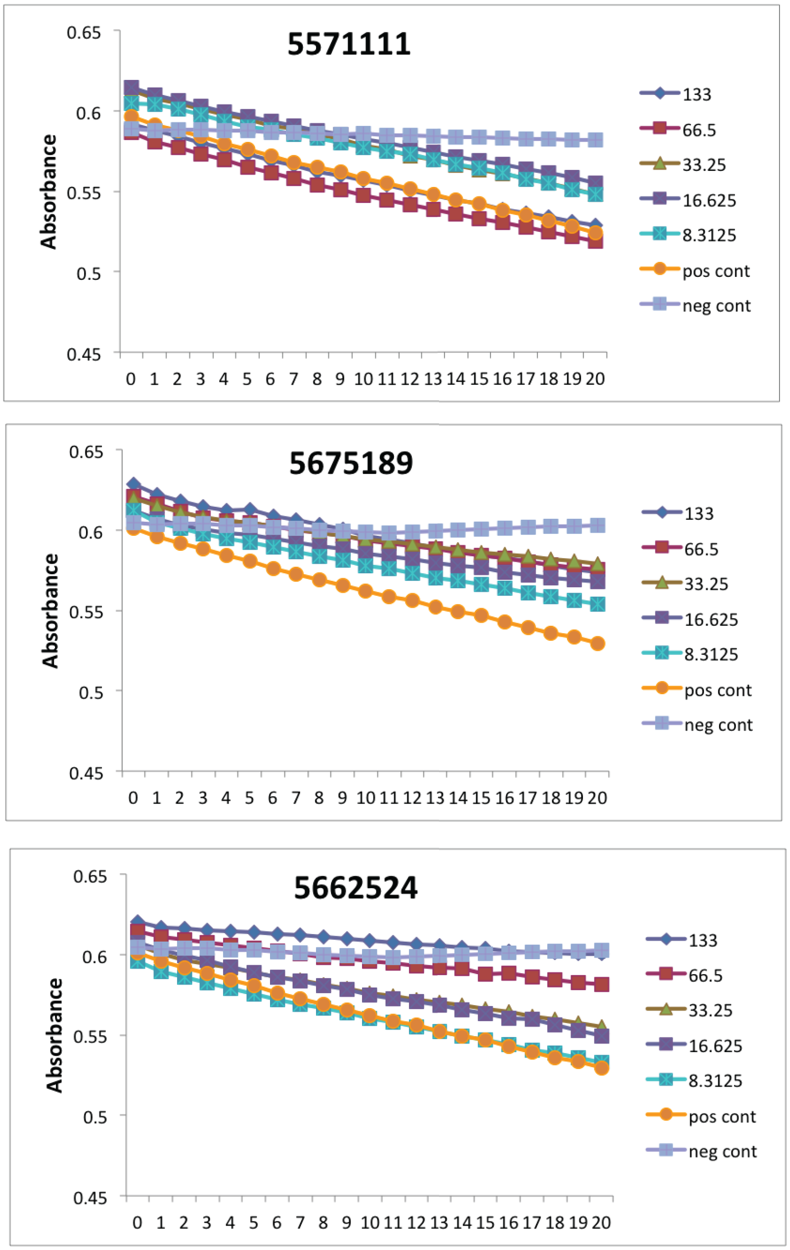


Figure 16. Secondary small molecule screen with αKG mimetics.
Panel displays three small molecules identified from initial high-throughput screen and advanced to the secondary screen. All three compounds had an inhibitory effect on mutant IDH1. The legend lists the concentrations (µM) used for the small molecules in the screen with the absorbance on the y-axis and time in minutes on the x-axis.

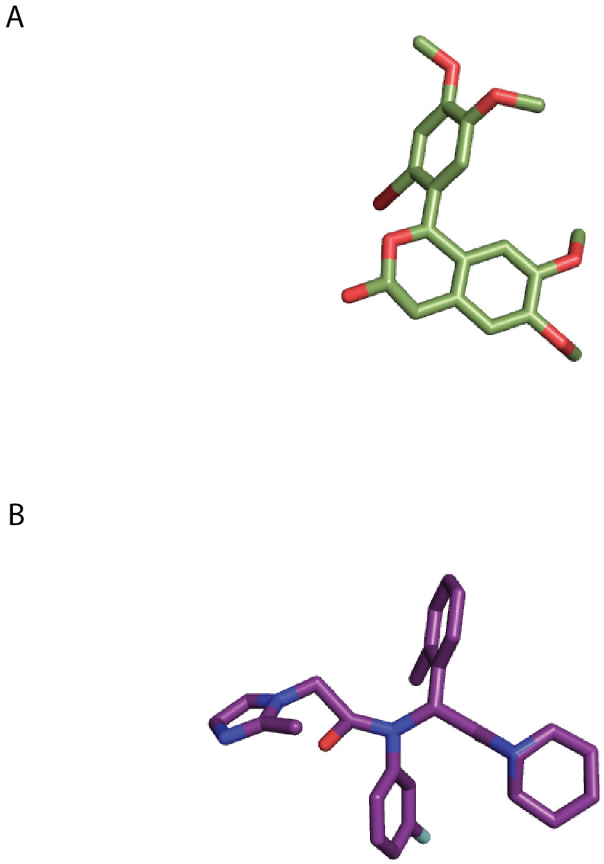


Figure 17. Chemical structures of small molecules used in studies.
(A) A07 and (B) ZAG were used in the NADPH depletion assays as well as in the structural determination of IDH1 R132H. A07 was identified during a high-throughput screen using a subset from a ChemBridge Library and ZAG is an IDH1 inhibitor identified by Agios Pharmaceuticals.

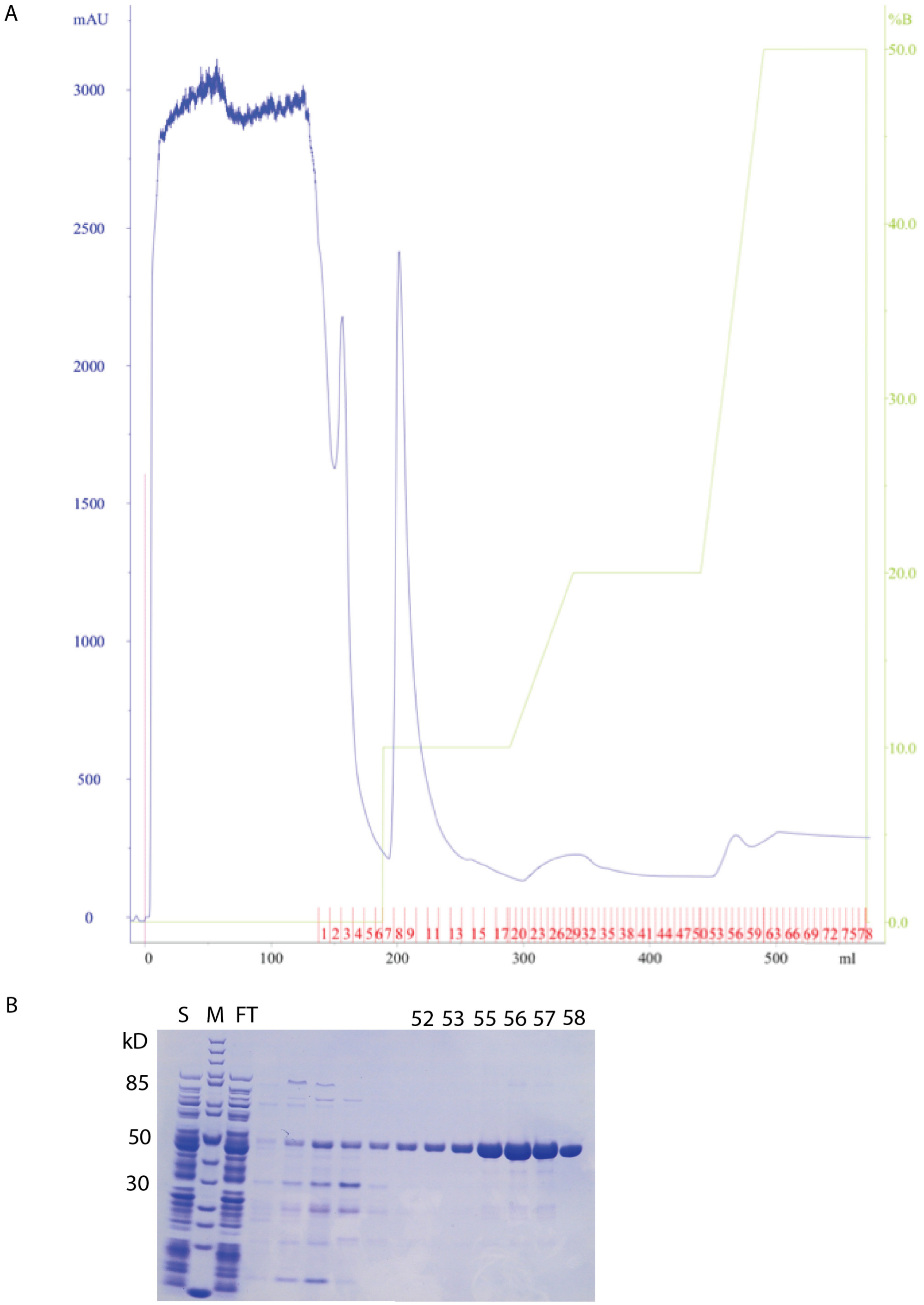


Figure 18. Ni^{2+}-affinity chromatography of Se-Met replaced IDH1 R132H.
(A) Chromatogram of purification. Blue line denotes absorbance at 280 nm in milliabsorbance units and green line represents gradient of elution buffer with concentration listed on right axis. Pink line denotes sample injection. (B) SDS-PAGE gel of purification stained with Coomassie blue stain. M refers to molecular weight ladder in kilodaltons. Relevant fractions are shown: S refers to soluble lysate, FT to unbound proteins or “flow through,” also fractions from 3rd peak in blue trace are shown.

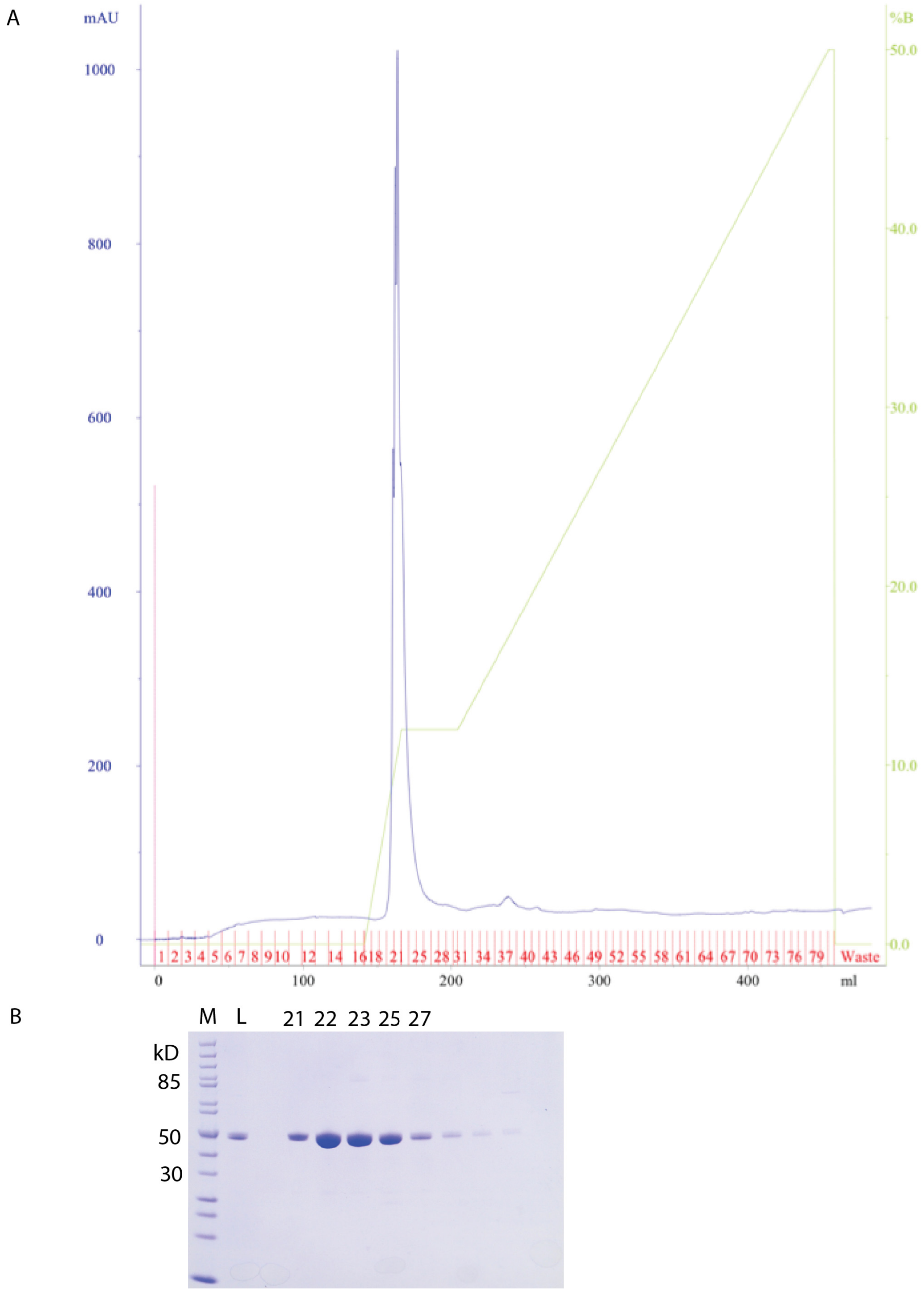


Figure 19. Anion exchange chromatography of Se-Met replaced IDH1 R132H.
(A) Chromatogram of purification. Blue line denotes absorbance at 280 nm in milliabsorbance units and green line represents gradient of buffer used for elution with concentration. Pink line denotes sample injection. (B) SDS-PAGE gel of purification stained with Coomassie blue stain. M refers to molecular weight ladder in kilodaltons and L to sample loaded onto column. Fractions corresponding to peak are shown.

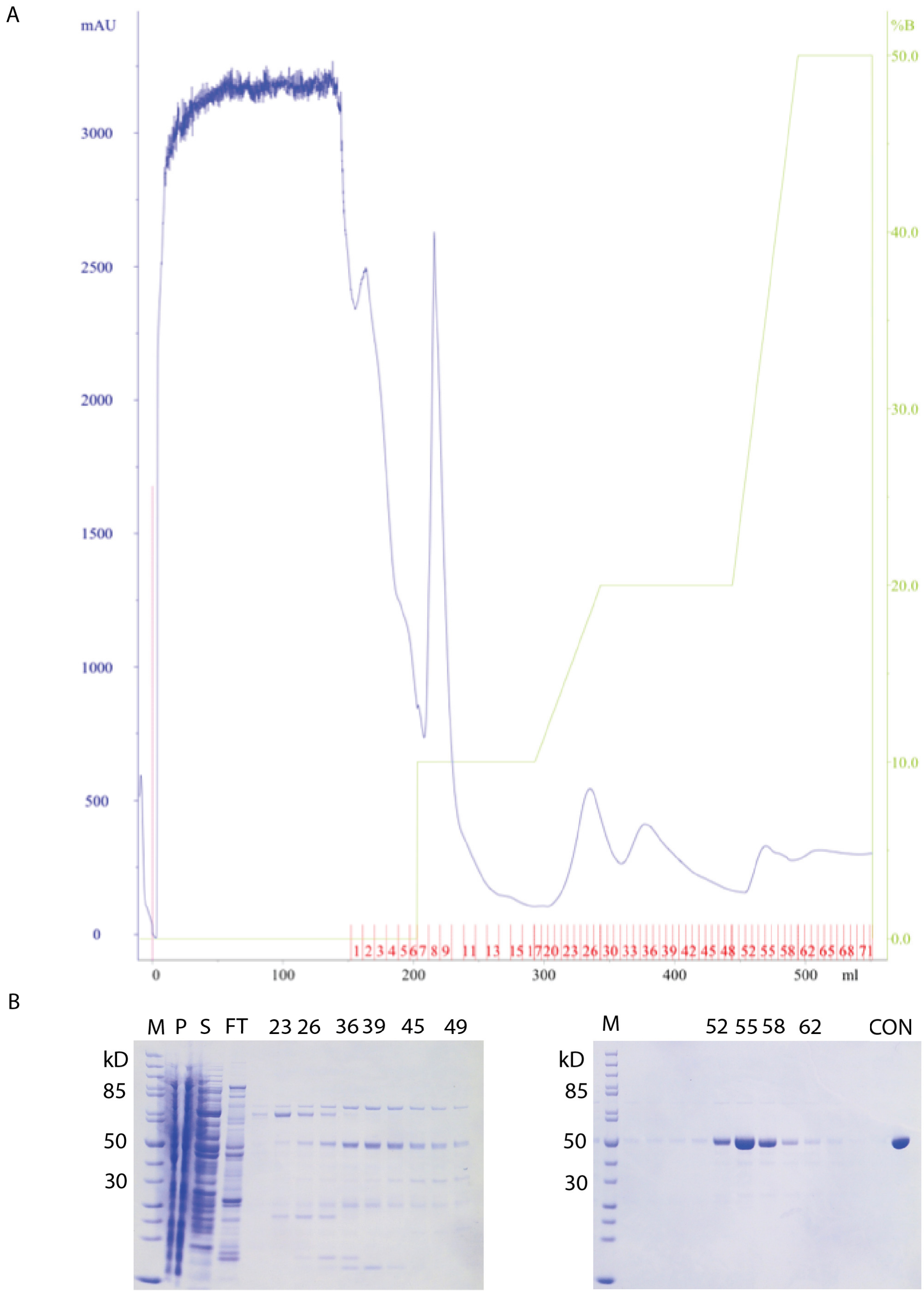


Figure 20. Ni^{2+}-affinity chromatography of IDH1 R132H.
(A) Chromatogram of purification. Blue line absorbance at 280 nm in milliabsorbance units and green line represents gradient of buffer used for elution. Pink line denotes sample injection. (B) SDS-PAGE gel of purification stained with Coomassie blue stain. M refers to molecular weight ladder in kilodaltons. P to insoluble pellet, S refers to soluble lysate, FT to unbound proteins or "flow through," fractions from 3rd peak in blue trace are shown to the right on 2nd gel, CON refers to control from previous purification.

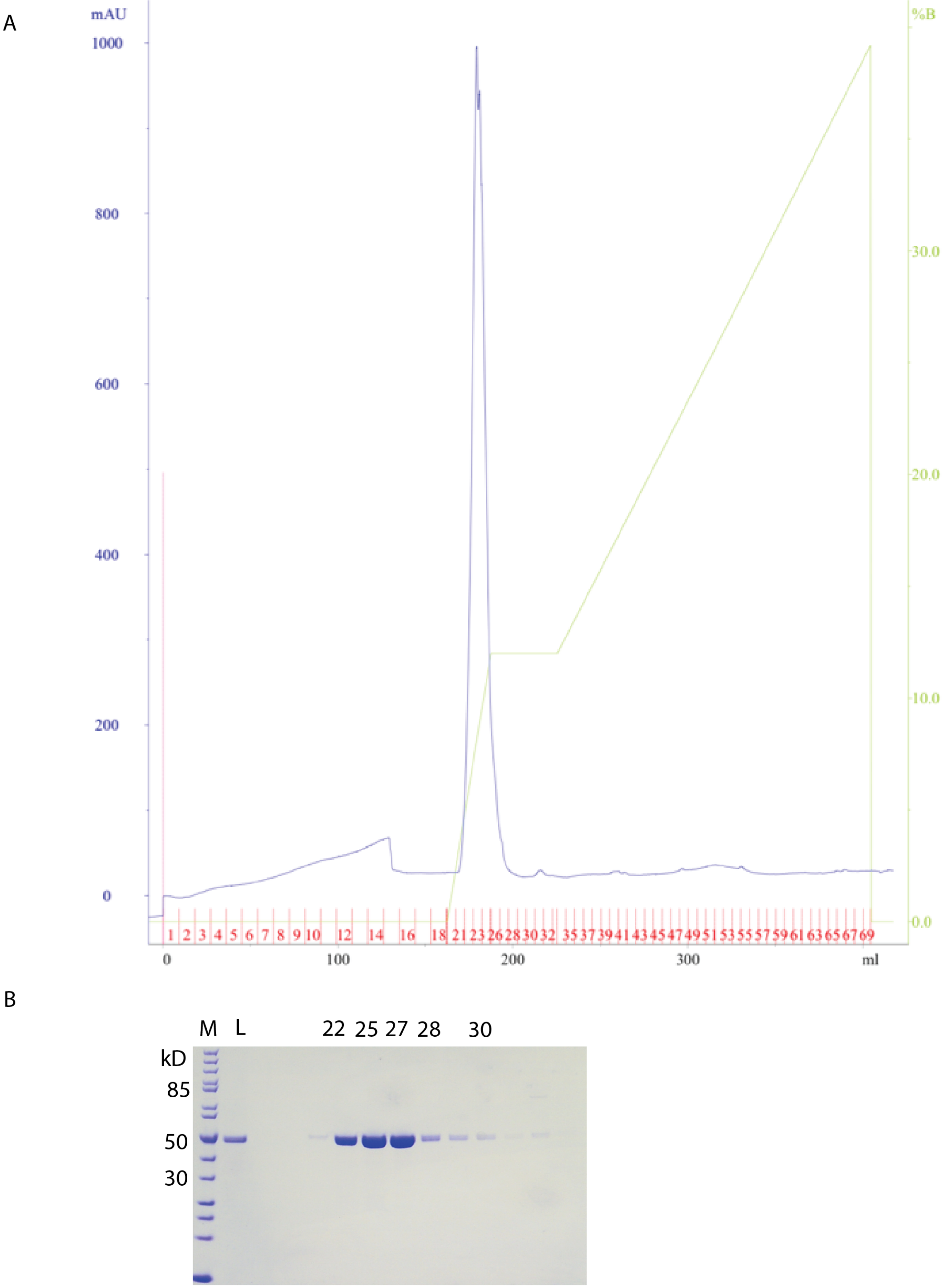


Figure 21. Anion exchange chromatography of IDH1 R132H.
(A) Chromatogram of purification. Blue line denotes absorbance at 280 nm in milliabsorbance units and green line represents gradient of buffer used for elution with concentration listed on right axis. Pink line denotes sample injection. (B) SDS-PAGE gel of purification stained with Coomassie blue stain. M refers to molecular weight ladder and L to sample loaded onto column. Fractions corresponding to peak are shown.

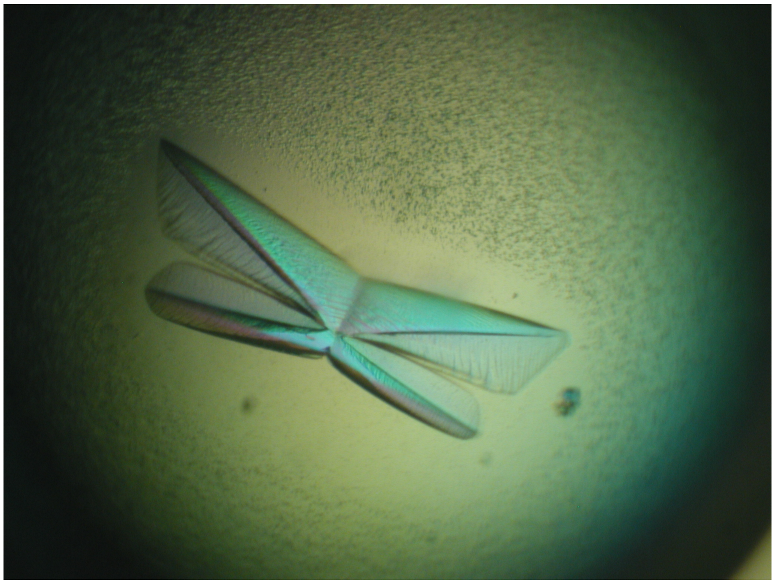


Figure 22. Crystal complex formed by IDH1 R132H, NADP$^+$, and ICT.
Crystals of IDH1 R132H were grown by hanging drop vapor diffusion. NADP$^+$co-purified with enzyme. Protein complexes were co-crystallized from solutions of protein mixed with isocitrate. Optimized crystals were tetrahedragonal in shape.

4. Mechanistic Insights from Structure Determination of IDH1 R132H

4.1 Activity of small molecules against IDH1 R132H

Three out of 48 selected compounds displayed IDH1 R132H inhibition in the primary screen at a concentration of 133 uM. Subsequently, these compounds were retested in oxidizing and reducing assays with ICT and $NADP^+$ vs. αKG and NADPH respectively, in a dose range from 8.3 uM to 133 uM. Three compounds inhibited IDH R132H in a dose dependent fashion in the reducing assay modality but no inhibition of IDH1 R132H was observed in the oxidative assay mode at any concentration. These compounds were advanced to a secondary screen and one compound, 5662524, (from this point on referred to as A07) was selected for biochemical assays in addition to ZAG. ZAG was identified from high-throughput screens conducted by Agios Pharmaceuticals. It has been shown to be an effective inhibitor against IDH1 R132H and reduces production of the oncometabolite 2HG in a glioma cell line [174].

Enzymatic inhibition assays were conducted with A07 and ZAG to further assess their inhibitory effects against IDH1 R132H. The consumption of NADPH was used to measure the activity of IDH1 R132H in the absence and presence of small molecules, as it converted α-ketoglutarate to 2-hydroxyglutarate. Substrate was added and all reactions were allowed to proceed for 4 hours. Reactions were monitored spectrophotometrically at 340 nm for the curve corresponding to the NADPH signal. Virtually all NADPH was consumed after 90 minutes in control reactions. However, in the presence of either 0.2 mM A07 (Figure 23) or 0.2 mM ZAG (Figure 24), IDH1 R132H was inhibited and the absorbance corresponding to the NADPH signal remained above zero. ZAG was more effective at inhibiting IDH1 R132H and blocking NADPH consumption than was A07. A07 limited the enzymatic activity of IDH1 R132H and reduced the consumption of NADPH. In contrast, the inhibition was more pronounced for reactions containing ZAG, the NADPH signal remained constant over the time course and there was virtually no change in the absorbance (Figure 24).

4.2 Structural insights from IDH1 R132H

4.2.1 Overall structural features

IDH1 R132H forms a homodimer in the asymmetric unit and belongs to the space group $P4_32_12$ (Table 4). Each monomer consists of 414 amino acids folded into three domains: large domain, small domain, and the clasp domain (Figure 25). Residues 1-103 and 286-414 make up the large domain, which has a Rossman fold typical of most oxidoreductases, while residues 104-136 and 186-285 make up the small domain which forms an α/β sandwich (Figure 26). A total of two-stacked anti-parallel β-sheets consisting of residues 137-185, form the clasp domain, which interlocks the two monomers. The small and clasp domains are the primary site of the interaction of the swapped domains. The region of the active site in both monomers is similar and yet not identical. As a reflection of the asymmetry of the monomers in the dimer, an RMSD of 1.42 Å is observed for 396 Cα atom pairs. The semi-open homodimer buries 2309 Å^2 upon dimer formation. In contrast, the closed homodimer buries 3600 Å^2 displaying a tighter interaction between monomers and the formation of a more compact dimer [175, 176].

Each monomer contains an active site cleft formed by the large and small domains. In the structures of IDH1 R132H-ICT-NADP$^+$ in the semi-open conformation, a molecule of NADP$^+$ binds at the NADP-binding site in each subunit, with isocitrate binding at 3.9 Å from the nicotinamide moiety and forming hydrogen bonds with amino acid residues of the large domain. Interestingly, NADP$^+$ is bound to the purified enzyme as reported for wild type IDH1, although none was added at the crystallization step. Similarly, in structures of IDH1 R132H in complex with αKG and NADP$^+$, a molecule of NADP$^+$ binds within each subunit at the NADP-binding site, and in this structure αKG is 3.8 Å from the nicotinamide moiety of NADP$^+$.

4.2.2 NADP$^+$ binds the semi-open conformation

The binding site of NADP$^+$ is formed by amino acid residues from the large domain of IDH1 R132H. The residues T77, H309, and G310 all form hydrogen bonds with NADP$^+$, in structures of IDH1 R132H in complex with αKG (Figure 27). In the structure of IDH1 R132H-ICT-NADP$^+$ the residues T75, N96 and E306 all form hydrogen bonds with the nicotinamide moiety of NADP$^+$, while the phosphates are recognized by the main chains of G310, T311, V312, and the side chain of R314. The residues T77, R82 and T311 recognize the ribose, while the adenine moiety of NADP$^+$ forms hydrogen bonds with N328 (Figure 28). Interestingly, a phosphate is consistently observed bound to R314 in all three structures. This recognition site differs from that of the closed structure (PDB code 3INM, [43]) where the adenine is recognized by V312, R324, H315 residues of one subunit and D253, N257 and K260 from the other. However, in both of our structures the adenine is only recognized by the N328 residue of one subunit.

4.2.3 Both α-ketoglutarate and isocitrate bind the semi-open conformation

In our structure, the α-ketoglutarate forms hydrogen bonds with water molecules, S94, R100, and is 3.8 Å from the nicotinamide moiety of $NADP^+$ in this conformation (Figure 29). The H132 that replaces an arginine residue at this position does not form hydrogen bonds with other residues in the active site. It is 7.9 Å away from the substrate. In the structure of IDH1 R132H-αKG-NADP-Ca, in the closed catalytically active conformation, the αKG interacts with T77, S94, R100, D275, from one subunit and T214 and D252 from the other subunit [18]. Both Y139 and K212 adopt different positions in the active site of our semi-open structure compared to those observed in the closed conformation.

In our structure of IDH1 R132H-ICT-$NADP^+$, the isocitrate is within hydrogen bonding distance to R100, S94, and T77, which differs from the isocitrate binding site in the closed conformation (Figure 30). This differs from the quaternary structure of the IDH1-NADP-ICT-Ca^{2+} (PDB code 1T0L, [18]), where R132 is within hydrogen-bonding distance to isocitrate and interacts with it. In that conformation, isocitrate also interacts with R109, R132, Y139 and D275 of one subunit, and K212, T214, and D252 of the other subunit. In this structure H132 is within 8.0 Å of the isocitrate.

4.2.4 Active site of semi-open conformation differs from closed conformation

Based upon substrate binding and release, IDHs undergo conformational changes that involve the rotation of the large domain relative to the small domain. In the closed conformation (PDB code 3INM, [43]) the large domain rotates 30° relative to the small and clasp domains (Figure 31). There are notable differences between the active site observed in the closed conformation in comparison to the semi-open conformation. The active site of the closed conformation is more compact. An RMSD of 4.71 Å is observed between the semi-open and the closed conformations (Figure 31).

In the wild type enzyme, residues 132-141 form a hinge that moves upon ICT binding, while residues 271-286 form a loop or 'regulatory segment' that protrudes into the ICT binding site but shifts substantially during catalysis [18]. However, in our structures, residues 271-286 of monomer A form a helix that is less flexible than the regulatory segment formed in the closed catalytically active conformation, and the equivalent loop or helix is missing in monomer B. Typically, a salt bridge is formed between R132 and the oxygen atom of D275 in subunit A; while in subunit B, R132 forms a salt bridge with the oxygen of D279. This aspartate residue, D279, also forms hydrogen bonds with the divalent ion allowing it to participate in catalysis. It is plausible that D252, D275, or D279 may be responsible for the initial proton abstraction from the hydroxyl of isocitrate, allowing the transfer of the hydride anion from the C2 of ICT to the C4 of the nicotinamide ring of $NADP^+$. The aspartate residues are also responsible for coordinating the divalent metal ion.

The arginine residues, R100, R109, and R132, are thought to promote proton abstraction by lowering the pKA of the metal-bound ICT. However, in our structures of mutant IDH1 in complex with either αKG or isocitrate, a histidine residue (H132) now replaces the arginine residue (R132) that would typically reside in this position so that D275 is now 6 Å from this H132, and D279 is 5.1 Å from this histidine residue.

Additionally, Y139 is not ordered, and cannot be traced from the binding site of the isocitrate in our structures of the semi-open conformation, but it is present in the active site of the closed conformation (Figure 32).

4.2.5 Small molecules bind mutant IDH1 at the dimer interface

To determine the binding site of the small molecules A07 and ZAG in IDH1 R132H, we determined the structure of IDH1 R132H-ICT-NADP-A07 by x-ray crystallography with data to 3.0 Å (Figure 33) and that of IDH1 R132H-ICT-NADP-ZAG to 3.3 Å (Figure 34). In both structures, the small molecules bind at the dimer interface of the semi-open IDH1 R132H (Figure 35). The complex structures IDH1 R132H-ICT-NADP-A07 and IDH1 R132H-ICT-NADP-ZAG possess great similarity to other semi-open structures (PDB code 3MAP, [177]) as shown by small RMSD of 1.2 Å between the Cα atom pairs upon superposition of monomer A. The allosteric binding site of A07 is between helix 252-259, loop 110-130 of monomer A, and opposite to the face of NADP of monomer B, 8 Å from the ribose and 26.8 Å from monomer A. The temperature factor of the binding site area is higher than that of the other portions of the structure, suggesting that the interface participates in the conformational changes required for catalysis. Both small molecules are approximately 8.6 Å away from IDH1 R132H. There was no binding observed of the small molecules to the wild type enzyme. These results suggest that both inhibitors bind in a pre-formed cavity that is not present in the closed structure (Figures 36 and 37).

The allosteric site observed in the inactive semi-open conformation of IDH1 R132H is asymmetric and locks the enzyme in a conformation that prevents the formation of the pre-transition state necessary for the enzyme to undergo catalysis

(Figure 35). In this conformation, the large domain is separated from the small and catalytic domain. Until now, only structures of IDH2 were determined in complex with inhibitors that displayed binding at the interface of the homodimer. This work shows that the preformed cavity allows for binding of IDH1 to allosteric inhibitors as well.

These structures confirm the dominant inhibition of the IDH1 mutation in the semi-open form and demonstrates that the dimer interface of IDH1 is druggable. It has a pre-defined pocket that participates in catalysis as it cycles between the open and closed conformations. Both compounds can be described as stabilizing and interfacial [161]. ZAG and AO7 inhibit the protein-protein interface required for catalytic activity and stabilize the semi-open conformation that precludes the conformation necessary for catalysis. As other 'stabilizers', they are also allosteric, binding at a site other than the active site. The inhibitors are interfacial and are juxtaposed to the interface using a grove present in one of the monomers. Using a combination of high throughput screening, inhibition assays, and structural biology, we identified a site of protein-protein interaction that provides an opportunity for selectivity different from the over-targeted sites of isocitrate and α-ketoglutarate. This work is yet another example of how structural biology is at the epicenter of the success of targeting protein-protein interactions.

Table 4. Data collection and refinement statistics for IDH1 R132H.

Crystal	**IDH1 R132H-NADP-αKG**	**IDH1 R132H -NADP-ICT-A07**	**IDH1 R132H-NADP-ICT-ZAG**
Space group	$p4_{3}2_{1}2$	$p4_{3}2_{1}2$	$p4_{3}2_{1}2$
Cell dimensions	a = 81.2 Å	a = 81.4Å	a = 80.9Å
	b = 81.2 Å	b = 81.4 Å	b = 80.9 Å
	c = 306.3 Å	c = 306.4 Å	c = 305.4 Å
	α=β=γ = 90.0°	α=β=γ = 90.0°	α=β=γ = 90.0°
Data Collection Statistics			
X-ray Source	BNL-X6A	BNL-X6A	APS-31
Wavelength (Å)	0.9793	0.9793	1.07
Resolution (Å)	50-2.1	50-3.0	50-3.3
(High Res shell)	(2.14-2.10)	(3.11-3.00)	(3.36-3.30)
Measured Reflections	640,895	766,097	730,587
Unique Reflections	61,099	19,076	16,176
I/σ	36.3 (3.5)	14.3 (4.1)	20.3 (6.8)
Completeness (%)	99.7	87.7(86.0)	99.3(100)
R_{symm} (%)	9.5	12.0	11.0
Refinement			
R_{cryst} (%)	19.9	21.9	19.7
R_{free}	23.5	30.8	27.6
Monomer in ASU	2	2	2
Protein atoms	7121	6979	6318
Water molecules	781	74	73
RMSD			
Bond length (Å)	0.015	0.013	0.019
Angle (°)	1.61	1.63	1.77
B-factor (Å^2)			
Protein	29.0	51.0	33.6
Ligand	29.0	68.0	69.6
Water molecules	33.4	30.5	22.6

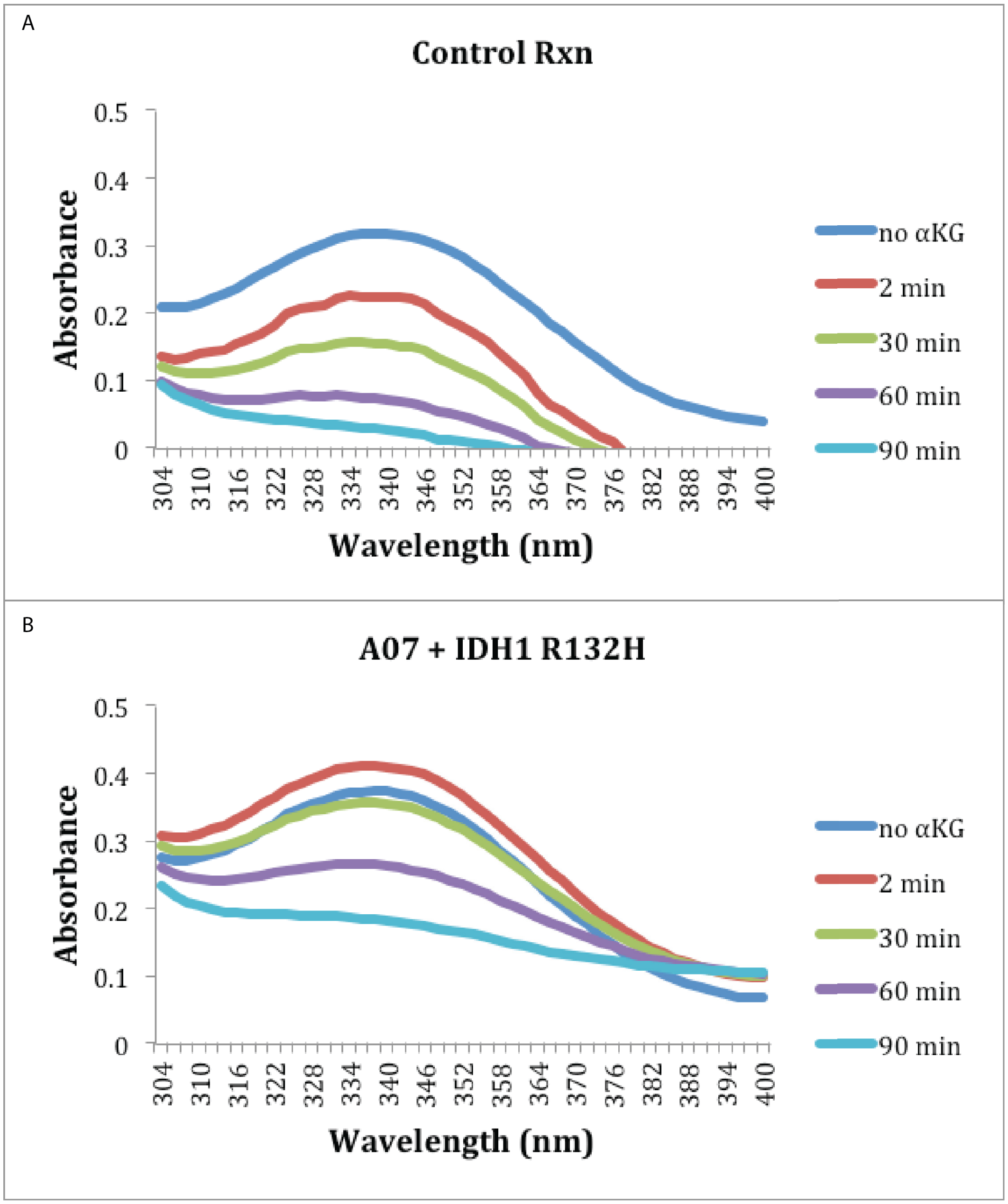


Figure 23. A07 blocks consumption of NADPH by IDH1 R132H.
(A) NADPH depletion assay performed in the absence of A07. (B) NADPH depletion assay in the presence of A07. Assays were followed spectrophotometrically at 340 nm to observe differences in the absorbance between reactions containing A07 and those without. The absorbance decreases as the reaction proceeds but, it does not drop to zero. This indicates that A07 limits the enzymatic activity of IDH1 R132H and blocks the consumption of NADPH. Scans were taken at the time points listed to the right in the legend.

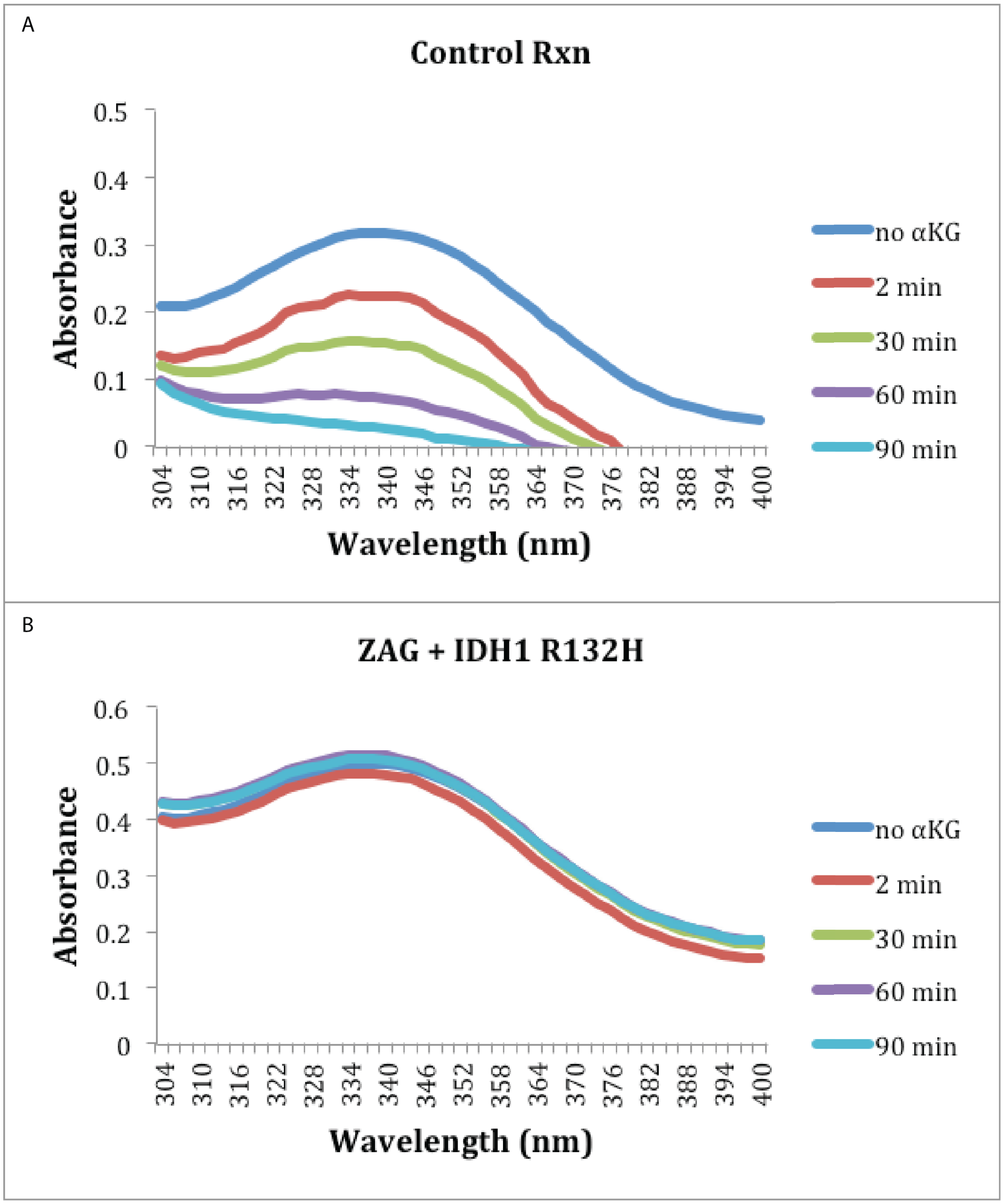


Figure 24. ZAG greatly reduces consumption of NADPH by IDH1 R132H.
(A) NADPH depletion assay performed in the absence of ZAG. (B) NADPH depletion assay in the presence of ZAG. Assays were followed spectrophotometrically at 340 nm to observe differences in the absorbance between reactions containing ZAG and those without. The absorbance remained constant as the reaction proceeded, indicating that ZAG inhibits IDH1 R132H and prevents the consumption of NADPH. Scans were taken at the time points listed to the right in the legend.

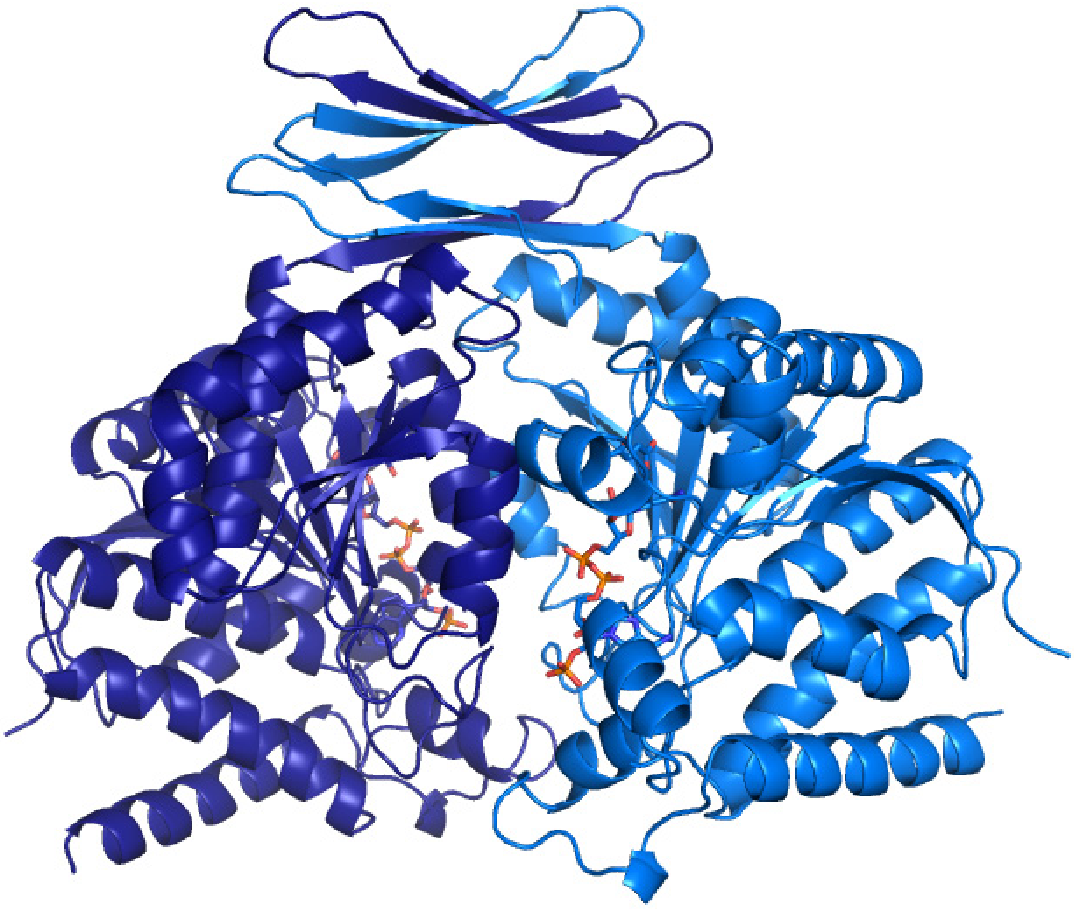


Figure 25. IDH1 R132H is an asymmetric homodimer.
Ribbon diagram displaying the overall structure of mutant IDH1 enzyme with individual chains shown in two different shades of blue. There is an active site present in each monomer. Within each active site, NADP$^+$ and α-ketoglutarate are colored by chain and shown as sticks.

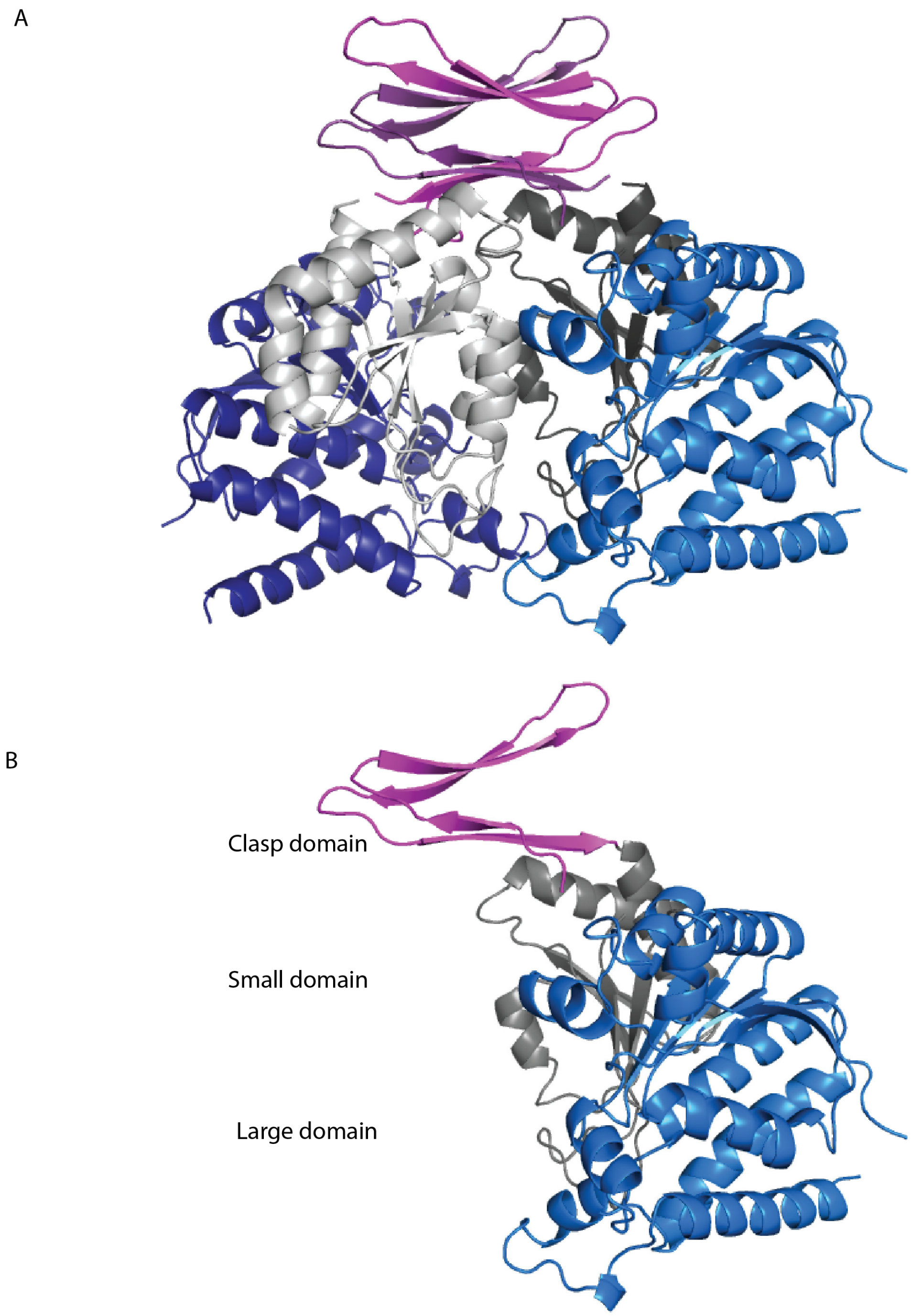


Figure 26. Monomers of IDH1 R132H have three domains.
(A) Ribbon diagram displaying overall structure of homodimer with both monomers shown. (B) Ribbon diagram depicting monomer A. Domains are colored in the following manner: the clasp domain is shown in magenta; the small domain is shown in gray; and the large domain is shown in blue. The clasp domain folds as two stacked anti-parallel β-sheets that interlock the two monomers. The small domain forms an α/β sandwich and the large domain has a typical Rossman fold.

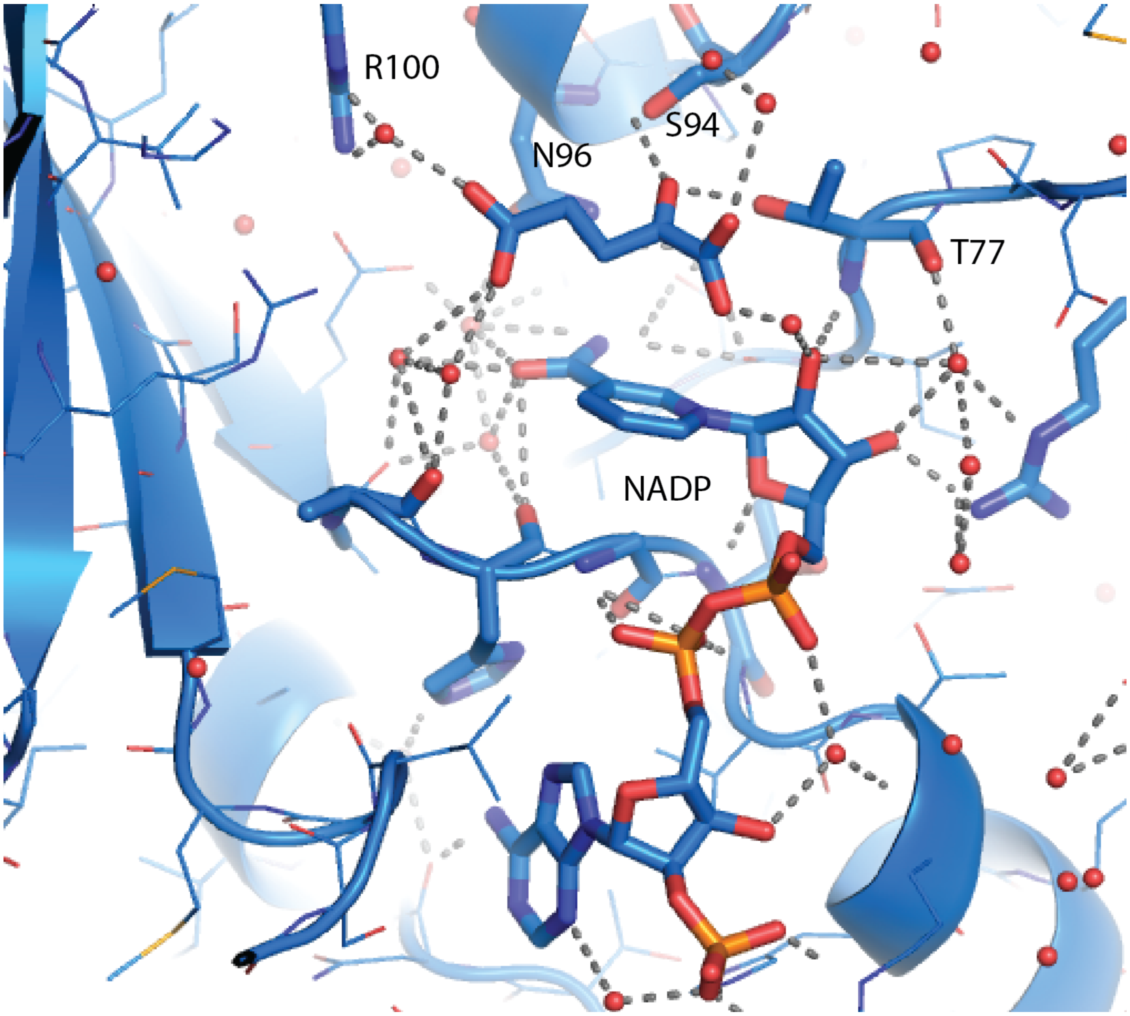


Figure 27. NADP$^+$ binds the semi-open conformation of IDH1 R132H.
A ribbon diagram displaying the NADP$^+$ binding site. NADP$^+$ interacts with amino acid residues of the large domain and the hydrogen-bonding pattern is displayed with gray dash marks. NADP$^+$ and amino acid residues are shown as sticks with carbons colored according to chain while water molecules are displayed as red spheres.

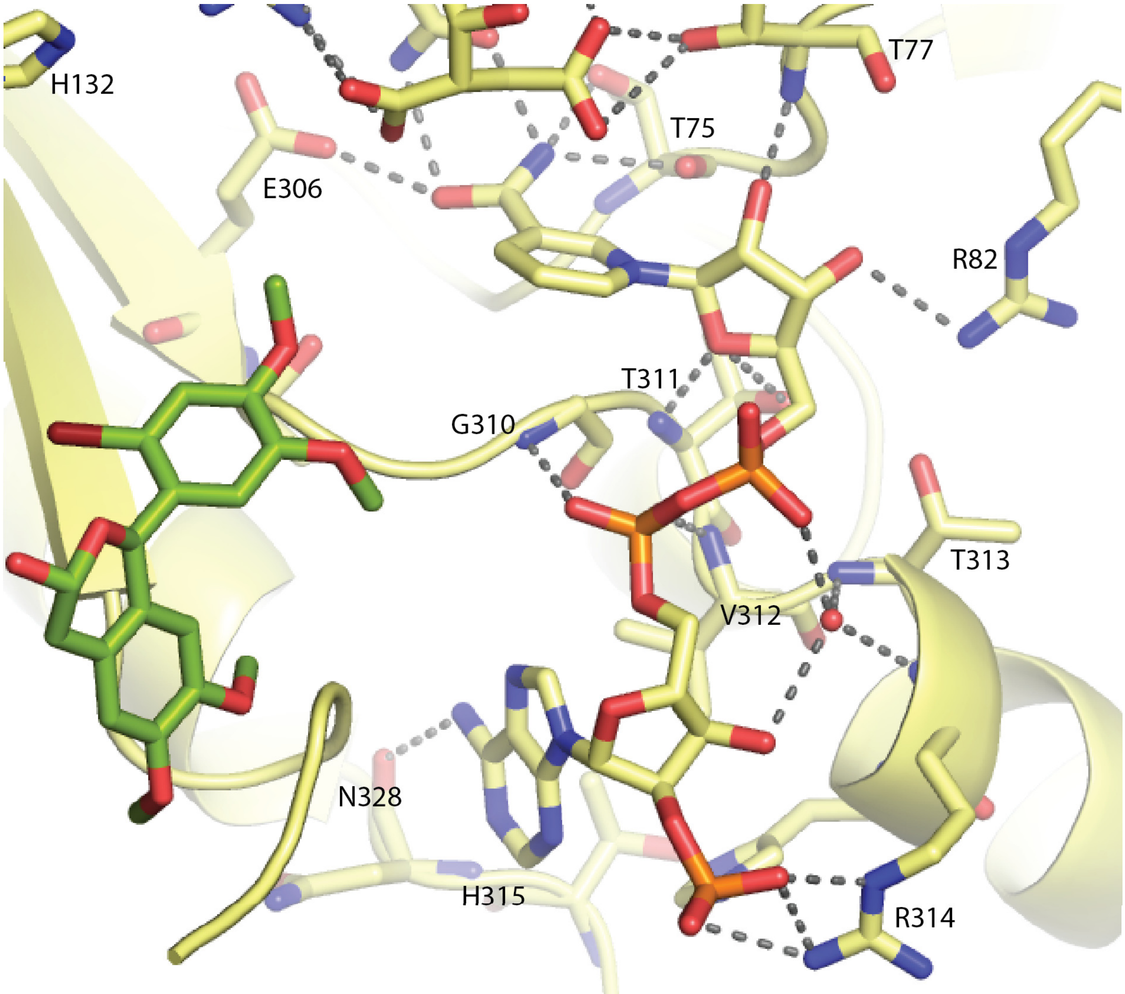


Figure 28. IDH1 R132H forms a complex with ICT, NADP$^+$, and A07.
Ribbon diagram of the active site showing the relationship between A07, isocitrate, and NADP$^+$. The binding site of NADP$^+$ is shown with interacting residues from the large domain. The hydrogen-bonding pattern is displayed with gray dash marks. Isocitrate, NADP$^+$, and amino acid residues are all displayed as sticks with carbons colored according to chain. Water molecules are displayed as red spheres and individual chains are shown in green and yellow.

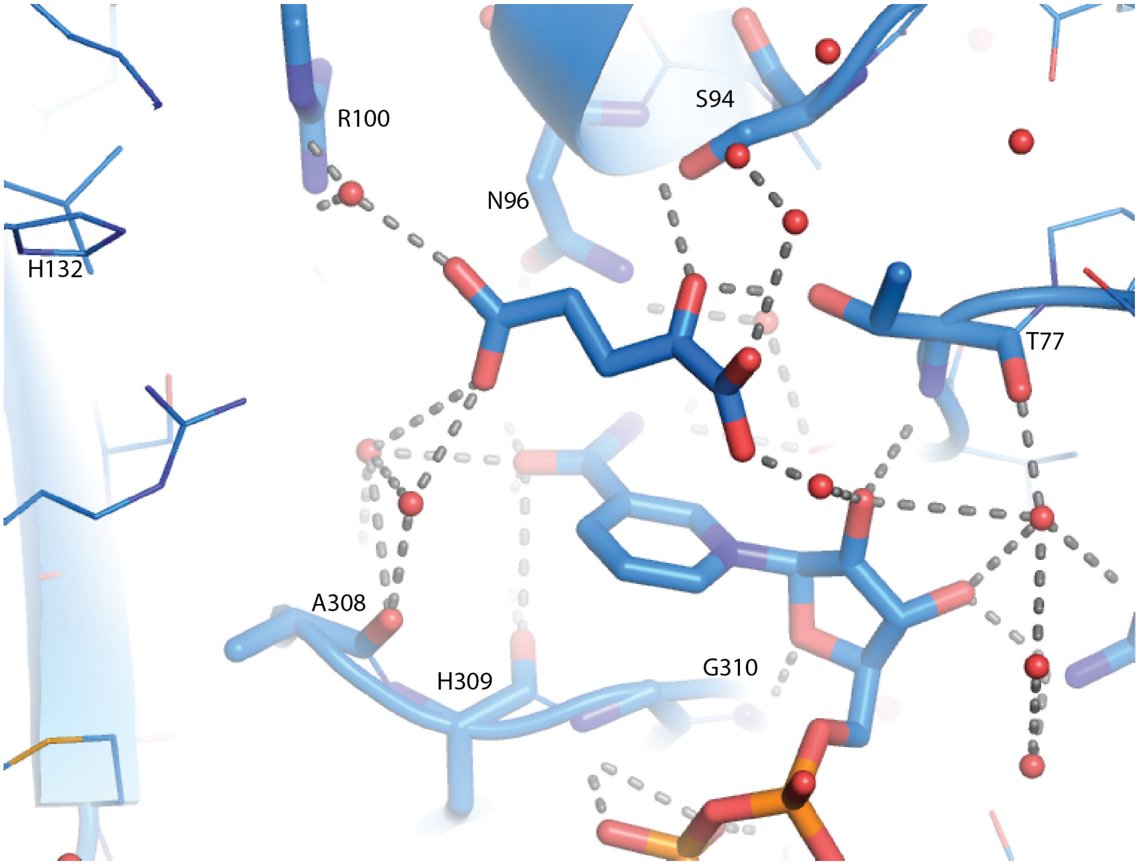


Figure 29. α-ketoglutarate binds semi-open conformation of IDH1 R132H.
Ribbon diagram displaying close-up of the active site showing the relationship between α-ketoglutarate and NADP$^+$ with individual chains shown in blue. The amino acid residues, α-ketoglutarate, and NADP$^+$, are all displayed as sticks with carbons colored according to chain and oxygen atoms shown in red. The hydrogen-bonding pattern is shown with gray dash marks and water molecules are displayed as red spheres.

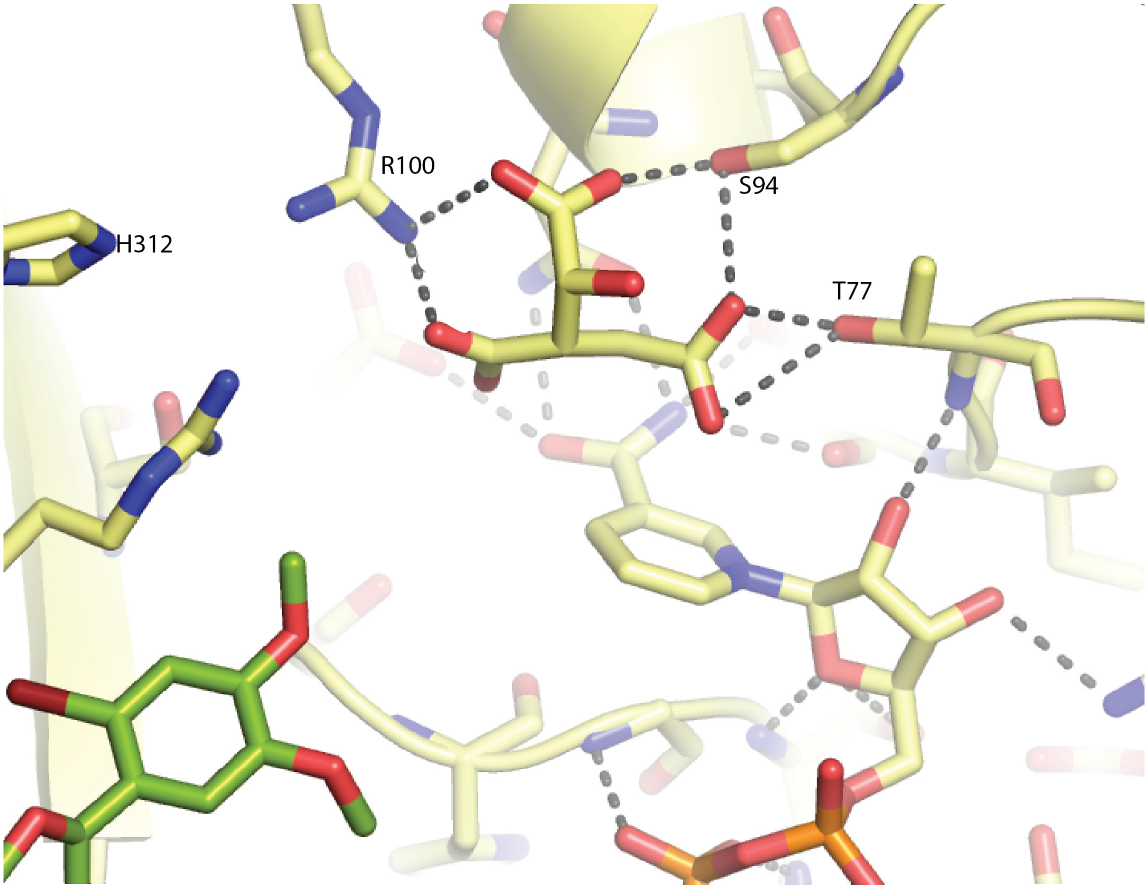


Figure 30. Isocitrate binds semi-open conformation of IDH1 R132H.
Close-up of the active site showing the relationship between A07, isocitrate and NADP$^+$ with individual chains shown in green and yellow. Isocitrate stacks over the nicotinamide moiety of the NADP$^+$. Isocitrate, NADP$^+$, and amino acid residues displayed as sticks with carbons colored according to chain. Hydrogen-bonding pattern is displayed as gray dash marks.

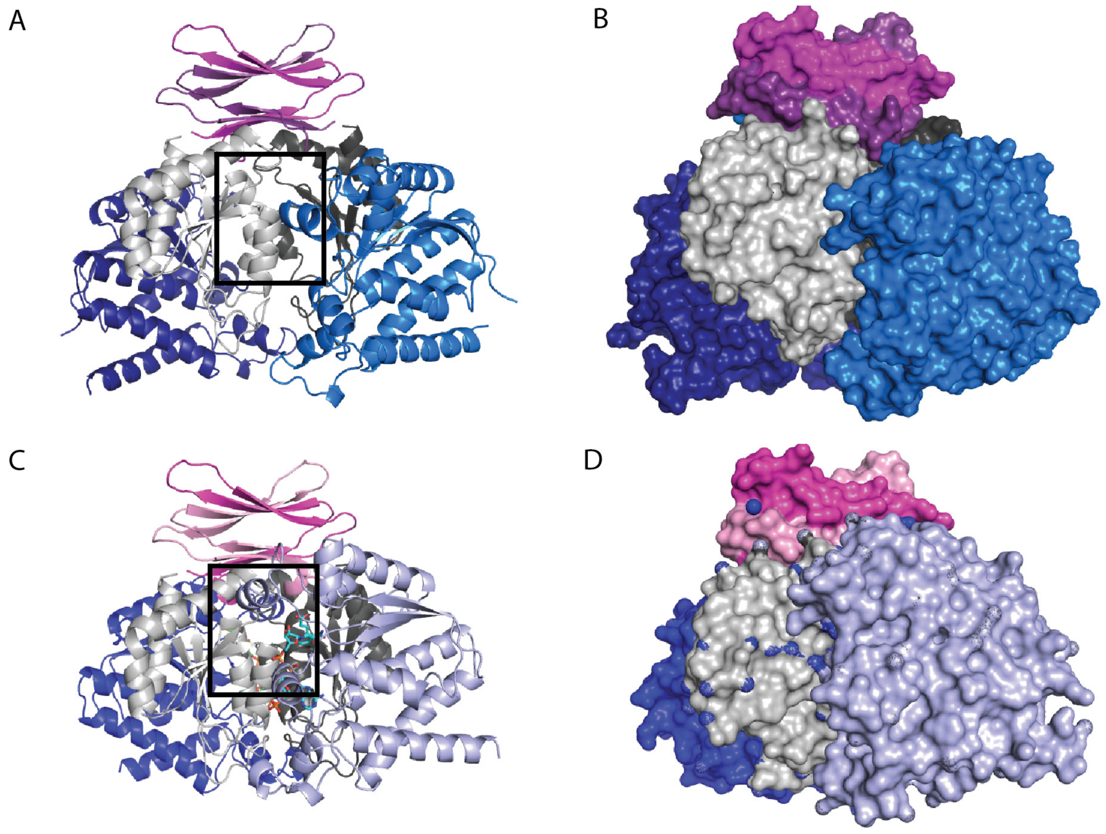


Figure 31. The active site of the closed conformation is more compact.
(A) Ribbon diagram of IDH1 R132H in the semi-open conformation. The clasp domain is displayed in two shades of pink, the small domain colored in two shades of gray, and the large domain is shown in two shades of blue. (B) Molecular surface representation of IDH1 R132H in semi-open conformation with domains colored as outlined in (B). (C) Ribbon diagram of IDH1 R132H in the closed conformation (PDB id: 3INM) with α-ketoglutarate and $NADP^+$ in the active site. The clasp domain is displayed in two shades of pink, the small domain is colored in two shades of gray and the large domain is shown in either blue or lilac. Within the rectangle, note the appearance of two helices present in the closed conformation that are not present in the semi-open conformation. The large domain rotates 30° relative to the small domain. (D) Molecular surface representation of closed conformation (PDB id: 3INM) with domains colored as outlined in (C).

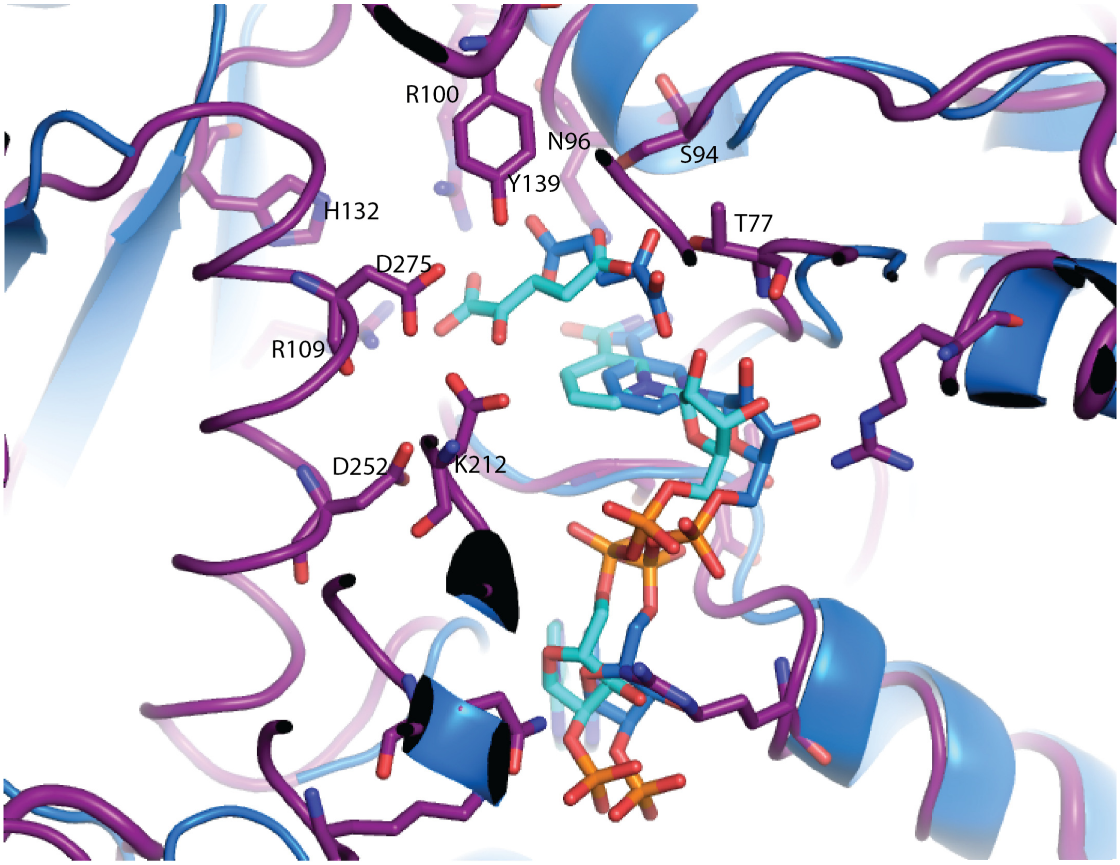


Figure 32. Active site of semi-open conformation differs from closed.
Ribbon diagram of the active site of the closed conformation superimposed over the active site of the semi-open conformation. The closed conformation (PDB id 3INM) is displayed in purple while the semi-open conformation is shown in blue. Residues from the active site of the closed conformation are colored in purple according to chain and include residues important for catalysis. NADP$^+$ and α-ketoglutarate from the closed conformation are displayed in purple with carbons colored according to respective chain. Note that NADP$^+$ and α-ketoglutarate have shifted position in the semi-open conformation. Both are displayed in blue with carbons colored according to chain.

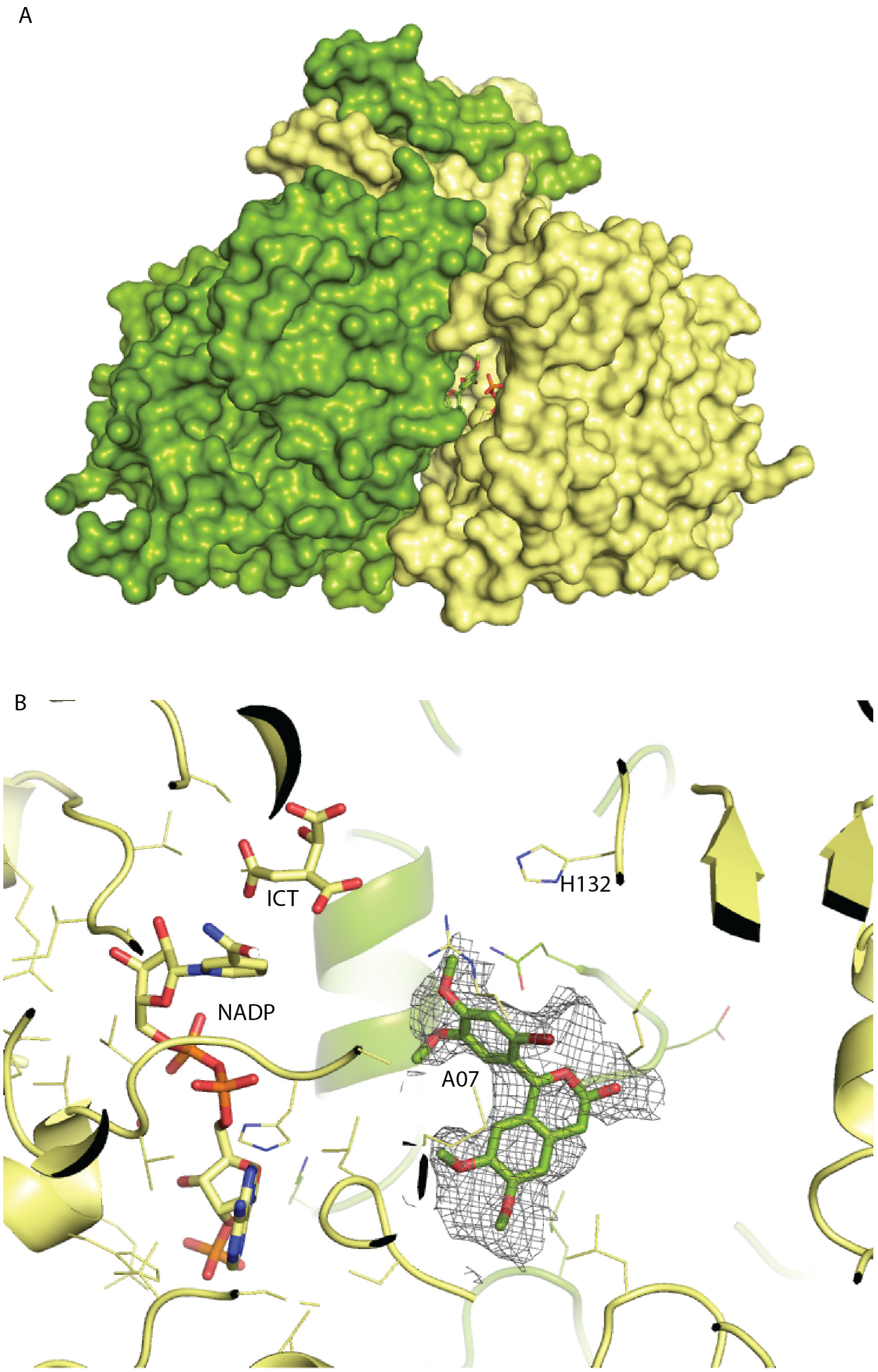


Figure 33. A07 binds mutant IDH1 at allosteric site.
(A) Molecular surface representation of the asymmetric homodimer of IDH1 R132H with individual monomers shown in green and yellow. A single molecule of A07 binds at the interface between the two dimers. (B) Ribbon diagram with close-up of binding sites of isocitrate, A07 with electron density map, and NADP$^+$. A07, isocitrate, and NADP all shown as sticks with carbons colored according to their respective chain.

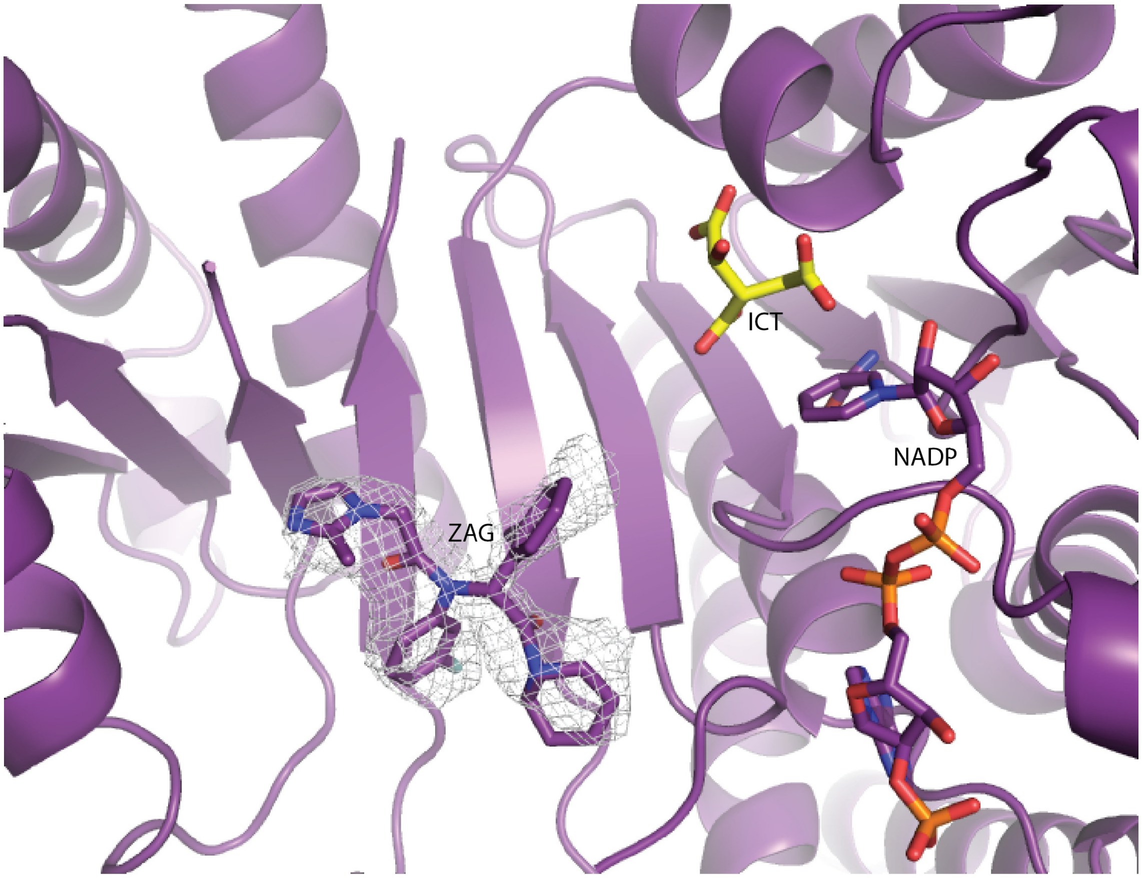


Figure 34. ZAG binds mutant IDH1 at allosteric site.
Ribbon diagram with close-up of isocitrate, ZAG, and NADP$^+$, displaying the allosteric site of binding. ZAG is shown in purple with an electron density map. NADP$^+$, ZAG, and isocitrate are all shown as sticks with carbons colored according to their respective chain.

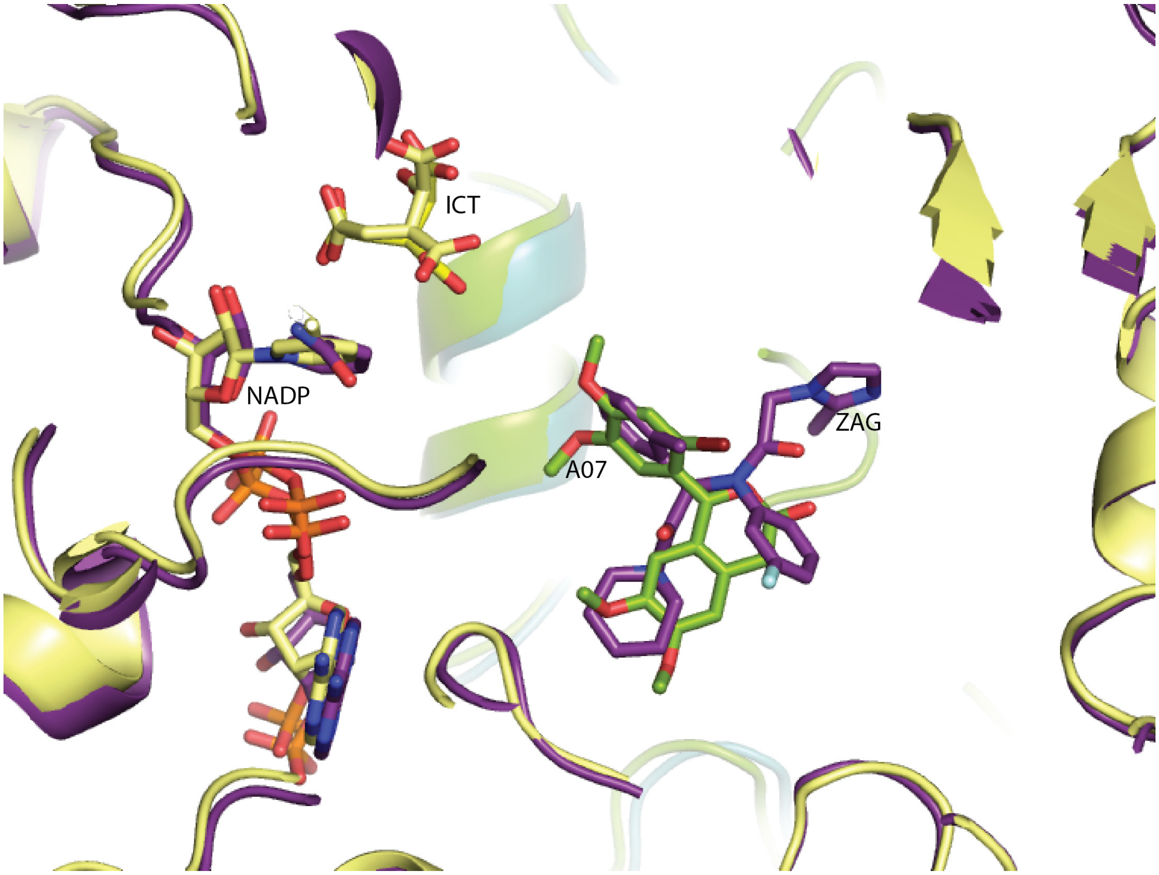


Figure 35. Small molecules bind at dimer interface of IDH1 R132H.
The binding site of A07 is superimposed over the binding site of ZAG. Both small molecules bind at the dimer interface. A07 is colored in green with individual chains shown in green and yellow ribbons. ZAG is colored in purple with respective chains displayed as cyan and purple ribbons. $NADP^+$ and isocitrate are colored according to the color of each individual chain.

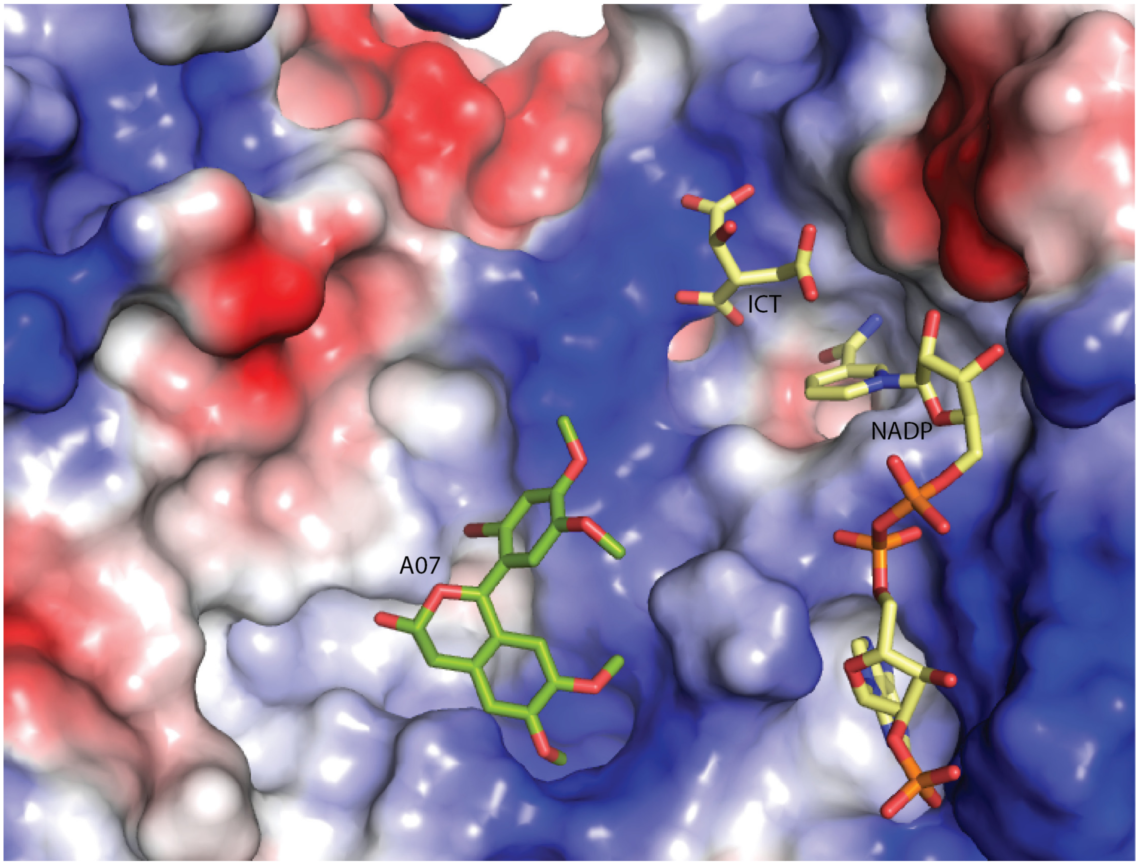


Figure 36. A07 binds mutant IDH1 in a pre-formed cavity.
A molecular surface representation of the dimer interface colored according to the electrostatic potential of the surrounding area. A07 binds in a pre-formed cavity and is displayed in green and shown as sticks. $NADP^+$, is shown with carbons in yellow and isocitrate is displayed with carbons in yellow, both are shown as sticks. Blue shading is for positive residues while red coloring is for negative residues.

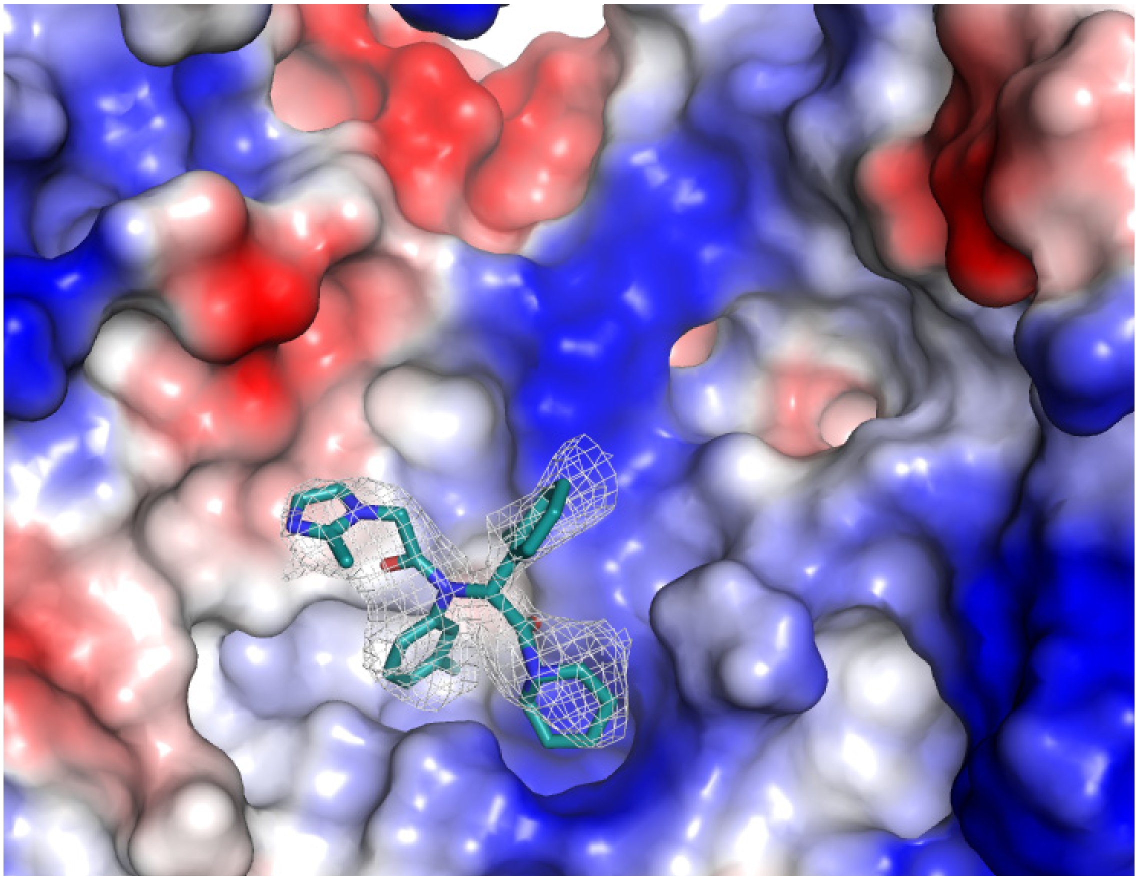


Figure 37. ZAG binds mutant IDH1 in a pre-formed cavity.
A molecular surface representation of the dimer interface colored according to the electrostatic potential of the surrounding area. ZAG is shown in teal and binds in a pre-formed cavity. Blue shading is for positive residues while red coloring is for negative residues.

5. Appendix I: Overview of the CREB Coactivator Complex

5.1 Introduction

Glucose is the main nutrient fuel of most organisms and is the sole energy source in some mammalian tissues and cells. The liver maintains blood glucose levels within a narrow range through the opposing actions of insulin and glucagon. In the fasting state, glucagon prevents hypoglycemia by stimulating glycogenolysis during the acute response and increasing gluconeogenesis during the chronic response. Postprandially, insulin suppresses both hepatic glycogenolysis and gluconeogenesis, thereby preventing hyperglycemia [178, 179].

Inappropriate glucose production is the major cause of the fasting hyperglycemia observed in individuals with type 2 diabetes mellitus (T2DM). Elevated glucagon levels together with impaired insulin signaling result in hyperglycemia and subsequent T2DM [180, 181]. Glucagon is considered a major etiological factor that worsens blood glucose control in patients with diabetes. Glucagon, acting through a G-protein coupled receptor, promotes the formation of a nuclear coactivator complex involved in the transcriptional regulation of hepatic glucose production. This complex consists of one transcription factor, cAMP response element binding protein (CREB), and two coactivators, CREB binding protein (CBP), and Transducer of Regulated CREB activity 2 (TORC2). Together, these proteins form the CREB coactivator complex (CREB/CBP/TORC2) [182, 183]. Insulin signaling results in the dissociation of the coactivator complex [182].

CREB belongs to the basic leucine zipper (bZIP) superfamily of transcription factors. It is homodimeric and possesses the following motifs: Q1 domain, kinase inducible domain (KID), Q2 domain, basic region/leucine zipper (bZIP) domain (Figure 38). The Q1 and Q2 domains are hydrophobic and glutamine-rich. Both of these domains interact with components of the basal transcription machinery and have demonstrated constitutive activation in binding studies with cultured cells [184, 185]. The KID domain is the region through which CREB interacts with CBP/p300, and includes a protein kinase A (PKA) phosphorylation site, as well as potential sites for other kinases [184, 186]. The bZIP domain is located towards the carboxy terminal, and the basic region mediates DNA binding, while the leucine zipper domain facilitates CREB dimerization [187, 188]. Within this family, CREB shares a high degree of similarity to the cAMP response element modulator (CREM) and the activating transcription factor 1 (ATF-1). These closely related proteins (Figure 39) comprise a subfamily referred to as the "CREB Family" of transcription factors [189-192].

Figure 38. Domain architecture of CREB.
The gray boxes represent Q1 and Q2. These are glutamine-rich domains and are also referred to as constitutive activation domains. The kinase inducible domain (KID) is displayed in red and contains a PKA phosphorylation site as well as other potential phosphorylation sites. The basic region/leucine zipper (bZIP) domain is shown in purple. The basic region mediates DNA binding, while the leucine zipper facilitates dimerization.

CREB is activated in response to many different stimuli and the putative list of target genes includes well over 100 genes. It has many binding partners and has been implicated in learning and memory pathways, as well as growth-factor-dependent cell survival, in addition to signaling cascades involved in glucose homeostasis. It binds *cis*-regulatory elements, also called cAMP response elements (CRE), located within the promoters of several different genes [193]. These CRE sites contain the palindromic consensus nucleotide sequence TGACGTCA [193, 194], although in some genes this

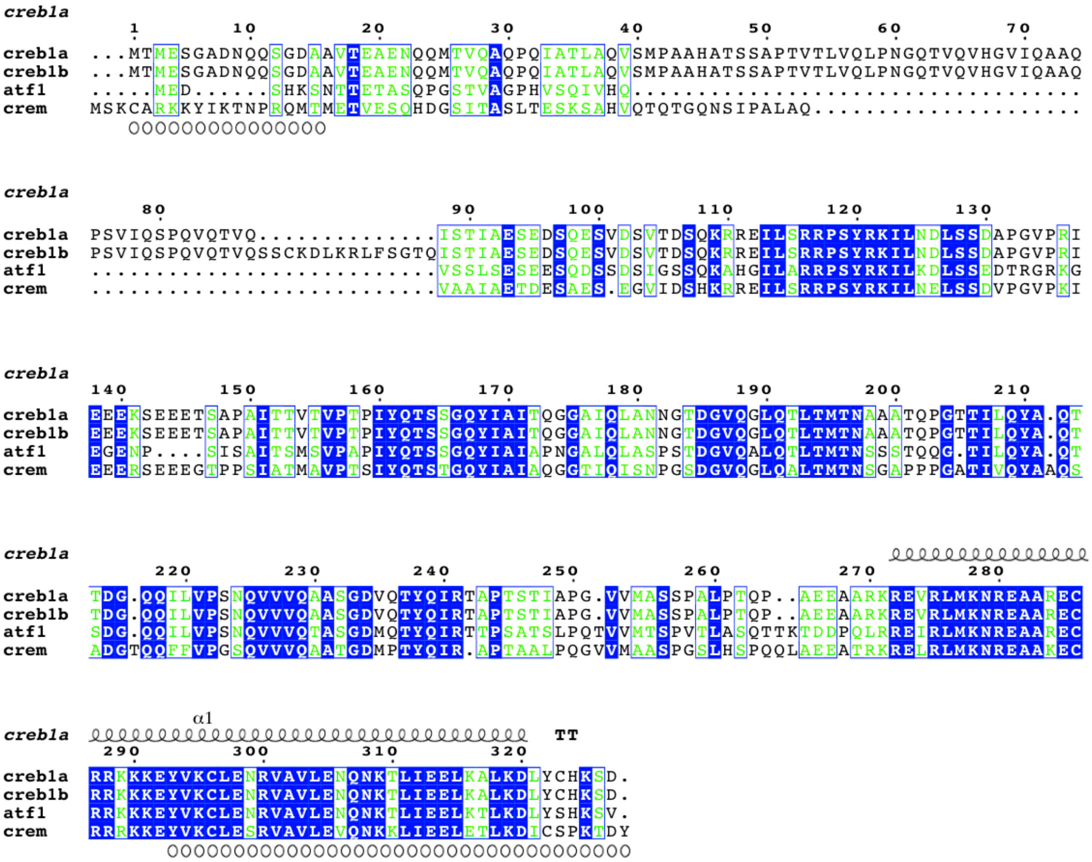


Figure 39. The bZIP family of proteins possess high sequence similarity.
CREB and CREM function as activators or repressors of transcription. They bind CRE sites located within gene promoters with high affinity and also form a subfamily which includes ATF-1. Conserved residues are displayed in blue and open circles denote deleted portions of the CREM construct used in our studies. Secondary structural elements for bZIP domain of CREB1A are displayed above the alignment.

motif is only present as a half site, CGTCA [195, 196]. There are two distinct classes of proteins known to mediate transcription of gluconeogenic genes by binding CREB: CBP/p300 family of transcriptional coactivators, and the TORC family of proteins. The phosphorylation of CREB is essential for cAMP-mediated activation of gene transcription and augments its activity by causing the recruitment of transcriptional coactivators such as CBP/p300 [197].

CBP and the paralogue p300 possess intrinsic histone acetyl transferase (HAT) activity, which allows them to modify histones and remodel chromatin. In this way, these coactivators can switch on transcription by making target gene promoters more accessible [198]. CBP/p300 have also been shown to be essential cofactors for several components of the basal transcriptional machinery [199, 200]. They play a critical role in the expression of a number of endocrine target genes and are closely related. Gene dosage of CBP/p300 is critical during development and homozygous CBP and p300 null mice are not viable [201]. CBP was first identified and characterized through its binding interactions with the KID of CREB [202]. Domain organization is as follows: nuclear receptor interacting domain (NR); cysteine/histidine rich region 1 (C/H1); kinase-inducible domain interacting domain (KIX); bromodomain (BROMO) binds acetylated lysine residues; histone acetyltransferase domain (HAT); cysteine/histidine rich region 3 (C/H3) containing both zinc finger and transcriptional-adaptor zinc finger (TAZ) domains (Figure 40). The NR is the domain that interacts with nuclear hormone receptors. The C/H1 and C/H3 are also referred to as TAZ1 and TAZ2 domains. They contain cysteine residues that participate in zinc-finger binding [203]. The KIX is the domain through

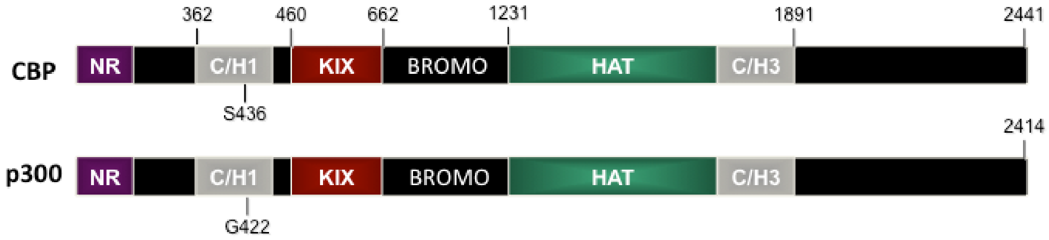


Figure 40. Domain architecture of the coactivators CBP and p300.
The nuclear receptor interacting domain (NR) can bind and interact with nuclear hormone receptors and is displayed in purple. The C/H1 (cysteine/histidine rich region 1) and C/H3 (cysteine/histidine rich region 3), contain transcriptional-adaptor zinc finger and zinc finger domains and are shown in gray. The kinase-inducible domain interacting domain (KIX) is the domain that interacts with CREB and is displayed in red. The bromodomain (BROMO) binds acetylated lysine residues and is shown in black. The histone acetyltransferase domain (HAT) is shown in green and catalyzes acetylation of histone proteins.

which CBP binds CREB, and it also serves as a scaffold or docking site for other transcription factors [204]. The bromodomain binds acetylated lysine residues and it enhances histone acetylation and promoter binding [205], while the HAT domain catalyzes the reversible acetylation of histone proteins present in nucleosomes [206].

Phosphorylation of CREB leads to the recruitment of the coactivators CBP/p300, which bind phospho-CREB through their KIX domain [200, 202]. A newer class of proteins referred to as TORCs, bind constitutively to the bZIP domain of CREB [207, 208]. Both CBP and TORC2 are transcriptional coactivators that have been reported to modulate glucose metabolism via the CREB coactivator complex [182, 183]. Given that inappropriate hepatic glucose production during fasting is the most important cause of fasting hyperglycemia, therapy has traditionally been directed at inhibiting this pathway.

5.1.1 Cyclic AMP signaling in the hepatocyte

The hepatocyte serves a critical physiological role in glucose homeostasis by storing or producing glucose based on the needs of the body. Binding of epinephrine, glucagon, and glucocorticoids to their respective G-protein coupled receptors will result in increased intracellular cAMP levels. In particular, following glucagon signaling, the cAMP signal will activate PKA, which phosphorylates the nuclear transcription factor CREB. The phosphorylation of CREB at S133 by PKA, leads to the recruitment of the nuclear coactivators CBP/p300 [209]. When CBP is phosphorylated it will not bind CREB. Activation of the cAMP pathway also leads to the dephosphorylation of TORC2. Following this event, TORC2 leaves the cytoplasm where it typically resides and enters the nucleus, where, like CBP, it functions as a transcriptional coactivator. As CREB, CBP and TORC2 (also referred to as CRTC2) assemble, glucose metabolism is modulated via the formation of this nuclear complex. Glucagon signaling increases hepatic glucose production by promoting the formation of this coactivator complex on CRE sites.

The PPARγ coactivator 1α (PGC-1α) regulates the gluconeogenic gene program in the liver. Transcription of *PGC1a* is mediated through cAMP signaling and it contains a well-defined CRE site to which the coactivator complex binds through CREB [182]. PGC-1α is a coactivator that binds associated transcription factors such as the hepatocyte nuclear factor 4 and the glucocorticoid receptor to regulate transcription of genes involved with glucose homeostasis [210]. Glucagon activates cAMP signaling and leads to an increase in the transcription of *PGC1a* and its downstream target genes. These target genes include those encoding the rate-limiting gluconeogenic enzymes phosphoenolpyruvate carboxykinase (PEPCK, *Pck1*) and glucose-6-phosphatase (G6P, *G6pc*).

5.1.2 Insulin signaling in the hepatocyte

In the fed state, insulin is a dominant suppressor of hepatic glycogenolysis and gluconeogenesis. Within the hepatocyte, insulin signaling stimulates the phosphorylation of CBP and reduces binding to CREB, regardless of the phosphorylation state of CREB. As such, insulin action can cause the termination or reduction of the cAMP signal mediated by CREB by preventing the binding of CBP to CREB, and causing the dissociation of the coactivator complex.

Previously, Wondisford and colleagues demonstrated that insulin phosphorylation at S436 is required for the efficient dissociation of the CREB co-activator complex and consequent attenuation of CREB-dependent gene transcription [182]. An *in vivo* mouse model of constitutively activated nuclear cAMP signaling was generated previously [211]. In this model, a germline mutation of CBP at S436 (S436A), removes a site of insulin-mediated phosphorylation. Without this site, CBP binds constitutively to phospho-CREB, and CREB-dependent gene transcription becomes resistant to the effects of insulin signaling. Insulin action in the hepatocyte would normally inhibit the binding of CBP to CREB and result in the dissociation of the coactivator complex. As a result of this mutation, these mice have impaired glucose tolerance due to the inappropriate activation of hepatic gluconeogenesis, hepatic insulin resistance, decreased insulin secretion from pancreatic β-cells, and increased β-cell mass [211].

There is a conserved protein kinase C (PKC) phosphorylation site in the C/H1 domain of CBP (Figure 41). This region contains the serine residue, S436, which lies in a location just N-terminal to the KIX domain on CBP, and it is this residue that is phosphorylated. Given its close proximity to the domain thought to interact with CREB, and the conservation of this serine residue among different species, it seems plausible that

S436 (m)
↓

Mus musculus	**428-VCLPLKNASDKRNQQ-442**
Homo sapiens	**429-VCLPLKNASDKRNQQ-443**
Macaca mulatta	**429-VCLPLKNASDKRNQQ-443**
Pan troglodytes	**429-VCLPLKNASDKRNQQ-443**
Rattus norvegicus	**428-VCLPLKNASDKRNQQ-442**
Equus caballus	**596-VCLPLKNASDKRNQQ-610**
Bos taurus	**603-VCLPLKNASDKRNQQ-617**
Gallus gallus	**418-VCLPLKNASDKRNQQ-432**
Xenopus laevis	**421-VCLPLKNASDKRNQN-435**

Figure 41. CBP phosphorylation site is conserved among eukaryotic species. Eukaryotic species are listed on the left. The residues surrounding the phosphorylation site are shown in black with the conserved serine residue displayed in red [182].

phosphorylation of CBP at this residue, could serve as a means of regulating hepatic glucose production. Through the dissociation of the coactivator complex, hepatic gluconeogenesis is reduced. They have also demonstrated that phosphorylation of CBP at S436 by insulin-mediated signaling, via the atypical protein kinase C ι/λ, is essential for dissociation of the coactivator complex (Figure 42).

In contrast, it has previously been suggested that the TORC (also referred to as CRTC) family of proteins are the predominant regulators of hepatic gluconeogenesis via CREB. Insulin signaling is thought to promote the phosphorylation TORC2, exclude it from the nucleus, and target it for proteasomal degradation (Figure 40). In this model of insulin regulation, S171 is phosphorylated on TORC2 leading to nuclear exclusion and the subsequent termination of CREB-dependent gene transcription [212]. The salt-inducible kinase 2 (SIK2) promotes the phosphorylation of TORC2 at S171 and inhibits its activity on the promoters of gluconeogenic genes [212]. Additionally, insulin signaling modulates the activity of TORC2 by acting through the serine/threonine kinase AKT, which phosphorylates SIK2 at S358. The insulin-mediated phosphorylation of TORC2 causes it to become sequestered in the cytoplasm where it is targeted for proteasomal degradation. However, in insulin-resistant states, TORC2 undergoes O-glycosylation and is retained in the nucleus [213].

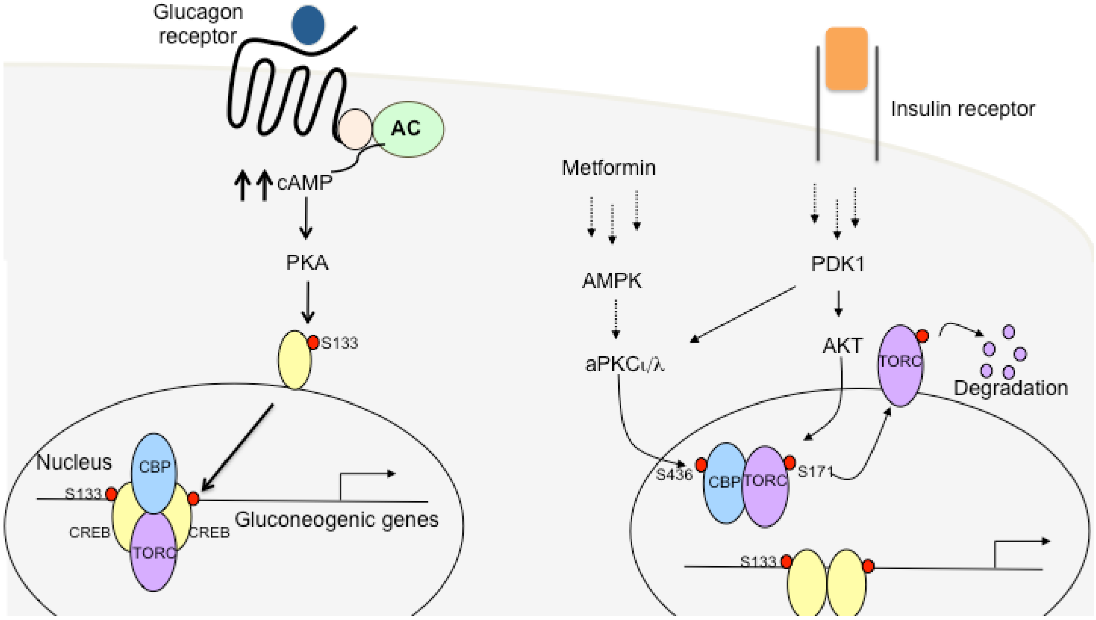


Figure 42. CBP plays central role in the regulation of hepatic glucose production. Glucagon signaling activates the assembly of the CREB coactivator complex. Insulin signaling caused the dissociation of this nuclear coactivator complex.

Recently, CRTC2/TORC2-deficient mice were derived to assess the role of TORC2 in hepatic glucose metabolism *in vivo.* After a prolonged fast, no significant difference was observed between the blood glucose levels of TORC2 mutants and littermate controls [214]. In addition, there was no difference in hepatic glucose production between these two groups of animals. Based on the findings of this group, TORC2 contributes to the transcriptional response but it is most likely not a key regulator of glucose homeostasis.

5.1.3 Cyclic AMP signaling in the adipocyte

The primary function of the adipocyte is to store triglycerides that are released as fatty acids and glycerol during lipolysis. The storage of excess energy as triglycerides is central to obesity and other metabolic disorders. Diet-induced thermogenesis is activated by the sympathethic nervous system (SNS) to increase energy expenditure and prevent obesity [215, 216]. The SNS regulates energy expenditure by sensing caloric excess. It innervates white adipose tissue (WAT) and brown adipose tissue (BAT) but, sympathetic innervation is thought to be more plentiful in BAT [217]. The binding of hormones and catecholamines to β-adrenergic receptors (βARs) located within the adipocyte will initiate a signaling cascade that stimulates lipolysis [215, 218]. βARs coupled to G proteins will mediate a cAMP signaling to activate PKA. Within this signaling cascade, PKA has downstream effects on other proteins.

Hormone sensitive lipase (HSL) and perilipin are both substrates of PKA, and their phosphorylation releases free fatty acids from triacylglycerol stored in lipid droplets within adipose tissue [219-222]. HSL hydrolyzes a variety of substrates including: triacylglycerol, diacylglycerol, and cholesterol esters. Perilipins help orchestrate the activity of HSL to facilitate lipolysis. They are located on the surface of the lipid droplet and inhibit the action of cytosolic lipases in the basal (unactivated/unphosphorylated) state [221]. A recent study conducted with obese women confirmed the protective effects of perilipin. In these women, perilipin protein content was decreased by 50% and basal and catecholamine-induced lipolysis was also increased by two- to four-fold [223].

Insulin terminates catecholamine-induced cAMP signaling through the degradative activity of phosphodiesterase 3, an enzyme that effectively reduces intracellular levels of cAMP. Hormones, cytokines, prostaglandins, nutritional inputs, gender, and physical activity all have an effect on lipolysis.

Alterations in cAMP signaling have the potential to impact whole body metabolism and more specifically lipolysis. Given the role of cAMP-mediated signaling in lipolysis, it would be interesting to observe the effects of the constitutive activation of this signaling pathway on energy balance. It is plausible that such signaling could result in a disruption of whole-body metabolism, or cause an increase in energy expenditure. Indirect calorimetry is a method that has existed for years but is being utilized as a research tool more and more to assess metabolic profiles. Of late, it is being used to investigate thermogenesis, as well as a number of disease pathologies implicated in metabolic disorders. It estimates energy expenditure and substrate utilization in a non-invasive manner through the use of gas exchange measurements. By employing indirect calorimetry, we can examine complex metabolic processes and correlate pathophysiological changes with interactions that occur on the molecular level. These studies will provide a better understanding of how diet-induced thermogenesis protects against obesity and determine how this mutation impacts metabolic rate.

5.2 Objectives

The overall goal of these studies was to define the structural and biophysical details of protein-protein interactions occurring between CREB and coactivators that result in the transcriptional activation of genes involved in hepatic gluconeogenesis and lipolysis. We sought to evaluate this altered signaling pathway within the adipocyte, and assess the resulting effects on energy balance using the CBP S436A mouse as an *in vivo* model of constitutively activated nuclear cAMP signaling. Additionally, we sought to provide a biophysical and biochemical basis for the assembly of the cAMP coactivator complex through the expression and purification of recombinant protein for biophysical and structural studies. The methods and results of the biophysical characterization of CREB, CREM and CBP will follow in subsequent appendices. This overview will also serve as an introduction for those appendices.

5.3 Sequence alignment of members of the basic leucine zipper family

The amino acid sequences of three members of the bZIP superfamily that together form the CREB subfamily of transcription factors, were selected for multiple sequence alignment. Protein sequences of two isoforms of CREB (CREB1A and CREB1B), ATF-1, and a truncated version of CREM, were aligned using Clustal Omega [162]. A figure containing the alignment was generated using the program ESPript [163].

5.4 Methods for metabolic phenotyping of CBP KI mice

5.4.1 Mouse Husbandry

All the animal procedures were performed in accordance with protocols approved by the Institutional Animal Care and Use Committee of the Johns Hopkins University. Animals were housed in a dedicated clean barrier animal facility within the Johns Hopkins University School of Medicine on a 12-hour light/12-hour dark photocycle. Mice had ad libitum access to food and water. Their diet consisted of either normal chow (Harlan Teklad) or a high fat diet deriving 60% of kilocalories from fat (Research Diets).

5.4.2 Genotyping and breeding of mouse cohort

A C57BL/6J mixed-background strain (originally obtained from Jackson Laboratories), bearing a heterozygous state (S436A/WT), were crossed to create both control and homozygous mice for this study. All mice generated for this study were from heterozygous pairings (S436A+/− × S436A+/−). The generation of the CBP S436A mutant mouse has been described previously [211]. Tail snips of up to 5 mm were obtained from mice at weaning, digested overnight at 55°C in buffer containing 10 mM Tris-Cl pH 8.0, 100 MM NaCl, 10 mM EDTA, 0.1 mg/mL Proteinase K, and 0.5% SDS. Subsequently, samples were processed to isolate genomic DNA, using either phenol/chloroform extraction followed by 2-propanol extraction, or the Gentra Puregene Mouse Tail Kit (Qiagen).

The S436A mutation was identified on genomic DNA obtained from tail snips using the following oligonucleotide primers: Common, 5′-GAC CTT GTT GCT TTG CAC TTG TTC-3′; Match, 5′-CTC CCT TTG AAA AAT GCC AG-3′; and Mismatch, 5′-CTC CCT TTG AAA AAT GCC GC-3′. PCR reactions proceeded for 32 cycles of denaturation at 95°C for 30 seconds, annealing at 61°C for 45 seconds, and extension at 72°C for 1 minute, with a final extension at 72°C for 7 minutes. All reactions were performed under standard conditions using 100 ng of genomic DNA, 7.5 nmol of primers, buffer containing 2.5 μl of 15 nM $MgCl_2$, and 0.5 μl of Taq polymerase per reaction. Reactions were run on 1% (w/v) agarose gel supplemented with ethidium bromide. Control mice displayed a band at the Match primer only, while CBP S436A mice only displayed a band at the Mismatch primer, and heterozygous mice (S436A+/−) displayed bands for both Match and Mismatch primers.

5.4.3 Assessment of body composition

The Metabolic Phenotyping Core Facility of the Johns Hopkins University School of Medicine performed the measurements. The EchoMRI-100 quantitative nuclear magnetic resonance (QNMR) system (Echo Medical Systems) was used to measure whole body composition parameters in CBP S436A mutant mice and age-matched littermate controls following consumption of normal chow or high fat diet. Direct measurements were taken *in vivo* of fat mass, lean mass, free water, and total body water. All measurements were performed in triplicate.

5.4.4 Indirect Calorimetry

Animals were separated and housed in individual cages 2 weeks prior to experimentation for adaptation. Only male mice were used for these studies. Homozygous CBP S436A mutants and control littermates were used for simultaneous assessments of daily changes in body weight, energy intake, and whole body metabolic profile in an open-flow indirect calorimeter. The Comprehensive Laboratory Animal Monitoring System (CLAMS, Columbus Instruments) was used to obtain these measurements as described below (Figure 43). Normal chow and HFD-fed mice were evaluated in separate studies. Mice were fed HFD for a total of 22 weeks during the course of these studies and metabolic assessments, with the exception of body composition studies, were made at the following time points following this challenge: 4 weeks, 8 weeks, 12 weeks, and 16 weeks.

Data were collected for 3 days to ensure that stable body weights and food intakes had been achieved, and to confirm acclimation to the calorimetry chambers. The data from the fourth day in the Oxymax chamber were analyzed. Rates of oxygen consumption (VO_2, ml·kg−1·h−1) and carbon dioxide production (VCO_2) were measured for each chamber every 14 min throughout the studies. Respiratory exchange ratio (RER = VCO_2/VO_2) was calculated by using CLAX Oxymax software (v. 4.02) to estimate relative oxidation of carbohydrate (RER = 1.0) versus lipid (RER approaching 0.7), not accounting for protein oxidation. Energy expenditure was calculated as $EE = VO_2 \times [3.815+ (1.232 \times RER)]$ [224]. Average metabolic values were calculated within subjects and then averaged across subjects for statistical analysis by Student's t-test, with $P \leq 0.05$, indicating significant group differences. Standard error was also calculated across groups.

5.5 Results and Discussion

CBP S436A mice were comparable in body weight to control mice at weaning and there was no obvious difference in length. However, on normal chow as time progressed, the mutants consistently weighed less than wild type littermate controls. Previously, it has been determined that CBP mutant mice have decreased body fat in comparison to wild type controls using QNMR. To probe the possible protective effects of this mutation on whole body metabolism, mice were challenged with high fat diet to induce metabolic stress. Mice were weaned onto normal chow and then switched to a high fat diet (HFD) at 8 weeks of age to observe differences in weight gain as well as other metabolic parameters.

At the start of these studies, before the introduction of HFD, control mice were approximately 16% heavier (29.2 ± 1.7 g) than CBP mutants (24.4 ± 1.6 g). Subsequently, 8 weeks after the introduction of HFD, wild type littermates on average gained more weight and were 20% heavier (38.8 ± 0.9 g) while homozygous mutants weighed less (31.7 ± 1.6 g) and accumulated less weight over this time period. Wild type littermate controls became obese while CBP S436A mutants maintained a leaner phenotype (Figure 44, A).

The differences in body weight between the homozygous mutants and littermate controls could be due to the accumulation of less fat mass. Not surprisingly, body composition analysis using quantitative NMR revealed that S436A mutants had less fat mass (14.1 g) relative to littermate controls (18.6 g) even after consumption of HFD for 22 weeks (Figure 44, B). This could account for the lower total body weight observed in these mice. Unfortunately, we did not harvest any fat pads to further assess the differences in fat mass between visceral or subcutaneous fat depots. Reductions in adipose tissue in response to HFD can be attributed to either an increase in physical activity, or a decrease in food intake. There were no observable differences in physical activity among CBP mutants versus control mice. However, CBP mutants experienced lower food intake in comparison with wild type mice, and this effect was observed within the first four weeks on the HFD. CBP S436A mutants were largely resistant to weight gain or diet-induced obesity (Figure 45), in comparison to controls.

Indirect calorimetry was performed to assess the effects of CBP S436A mutation on the metabolic rate. On regular chow, these analyses only revealed slight differences between controls and CBP mutants (data not shown). However on HFD, the differences between the two groups became more pronounced. To calculate RER and estimate energy expenditure, rates of oxygen consumption and carbon dioxide release were measured. Oxygen consumption is often used as a measure of basal metabolic rate. The CBP mutants displayed increased oxygen consumption (VO_2) and carbon dioxide production (VCO_2) relative to wild type controls. These results are indicative of an enhanced metabolism (Figure 46). Over time, both groups shifted from the use of carbohydrates as a fuel source with RER near 1.0, toward a preferential utilization and oxidation of lipids with RER approaching 0.7 (Figure 47). Although lower average RER are expected with consumption of HFD, compared to littermate controls, CBP mutants had a lower RER after 16 weeks on HFD, especially during the dark photocycle, (0.767 ± 0.014 versus 0.790 ± 0.015). CBP mutants experienced a greater shift toward enhanced whole-body oxidation of lipids, and increased energy expenditure despite hypophagia.

Although these results are promising, these findings need to be repeated and confirmed with larger cohorts ($N \geq 8$) to ensure reproducibility. The preliminary results of this work demonstrate that increased cAMP activation does lead to disruptions in whole body metabolism that could possibly protect against diet-induced obesity (DIO). For all of the metabolic parameters evaluated (body composition, food intake, oxygen consumption, RER, energy expenditure), there were subtle differences between homozygous mutants and littermate controls at baseline that became more apparent when challenged with HFD. Overall, the CBP S436A mutants were hypophagic, weighed less, had decreased fat mass, increased oxygen consumption, enhanced metabolic rate, enhanced energy expenditure, and displayed a metabolic profile consistent with a shift towards the preferential utilization of lipids as an energy source.

5.6 Future directions for S436A CBP mutant mouse

It was interesting to note that the S436A homozygous mutants seemed to experience appetite suppression and were hypophagic throughout the duration of the study. More than likely, this was a contributing factor to the lean phenotype maintained by these mutant mice. Meals rich in dietary fats also have the potential to initiate a greater postprandial response than would carbohydrates or proteins. Additionally, fasting levels of peptide YY (PYY) are elevated in several gastrointestinal diseases associated with loss of appetite [225, 226]. Following the consumption of a meal, hormones are released from the gastrointestinal tract, such as PYY. The anorexigenic POMC (pro-opiomelanocortin) neurons are activated by PYY, but it also simultaneously inhibits the orexigenic NPY/AGRP (neuropeptide Y/agouti related peptide) neurons, and regulates food intake by suppressing appetite within the arcuate nucleus [227].

The possibility remains that the CBP S436A mice have some defect in the hypothalamus or neural circuits that when coupled with the sustained activation of the cAMP pathway results in these anorexigenic effects on appetite and offers some protection against DIO. These mice might also secrete higher levels of PYY, and this could also lead to a loss of appetite and result in lower food intake. It would be interesting to follow up with experiments that evaluate serum levels of PYY and other gut hormones pre- and post-consumption of HFD to tease out what factor causes the hypophagia observed in these mice.

Figure 43. Oxymax open-flow indirect calorimetry chambers.
Fully automated open circuit indirect calorimeter used for metabolic studies with S436 CBP mice. Heat is derived by assessment of the exchange of oxygen for carbon dioxide that occurs during the metabolic process. The relationship between the volume of gas consumed (oxygen) and of that produced (carbon dioxide) reveals the energy content of the foodstuff utilized by the subject. This ‘calorific value’ is then applied to the volume of gases exchanged to compute heat and calculate energy expenditure.

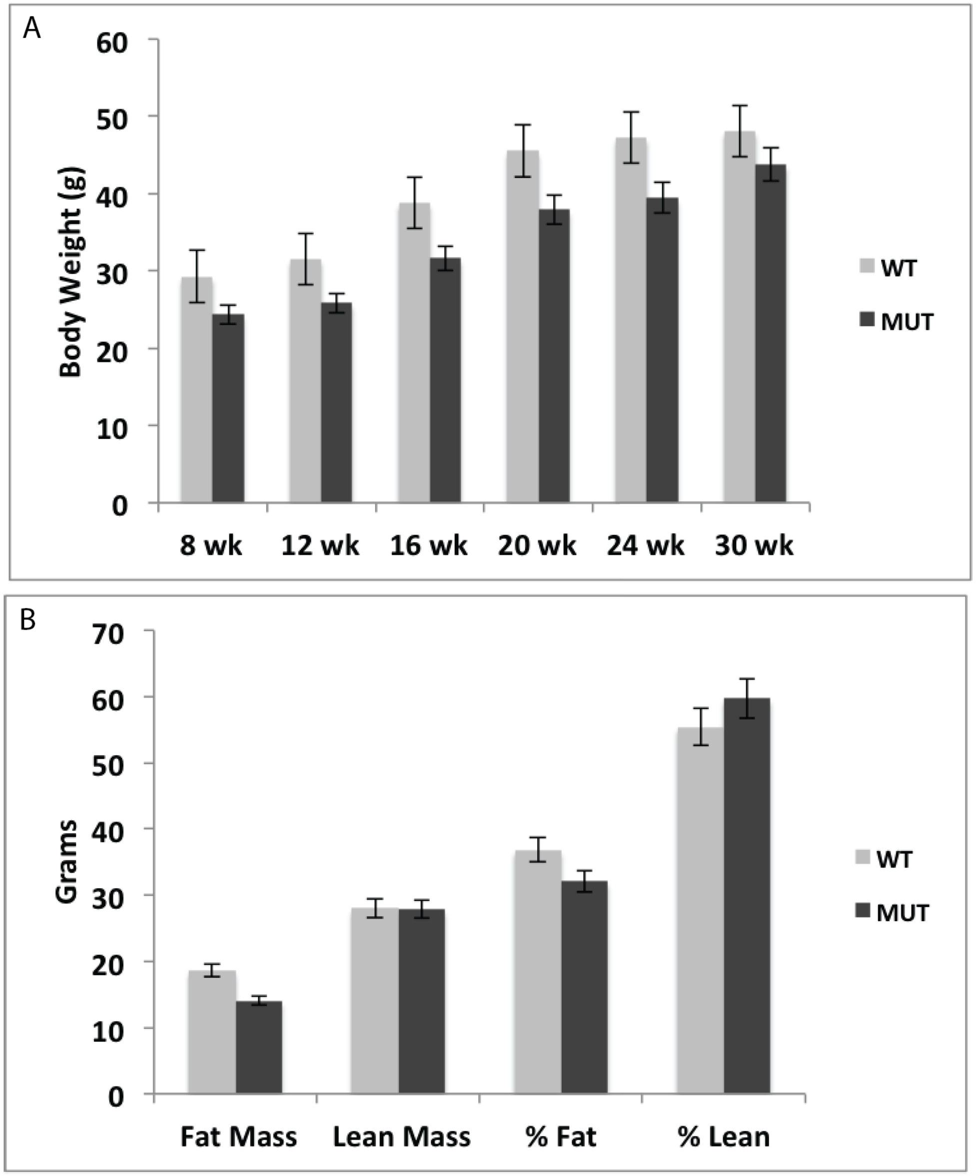


Figure 44. CBP mutants have a lean phenotype.
Effects of HFD on weight and body composition. A) Histogram of weight gain with number of weeks indicating age. B) Body composition analyses after 22 weeks on HFD.

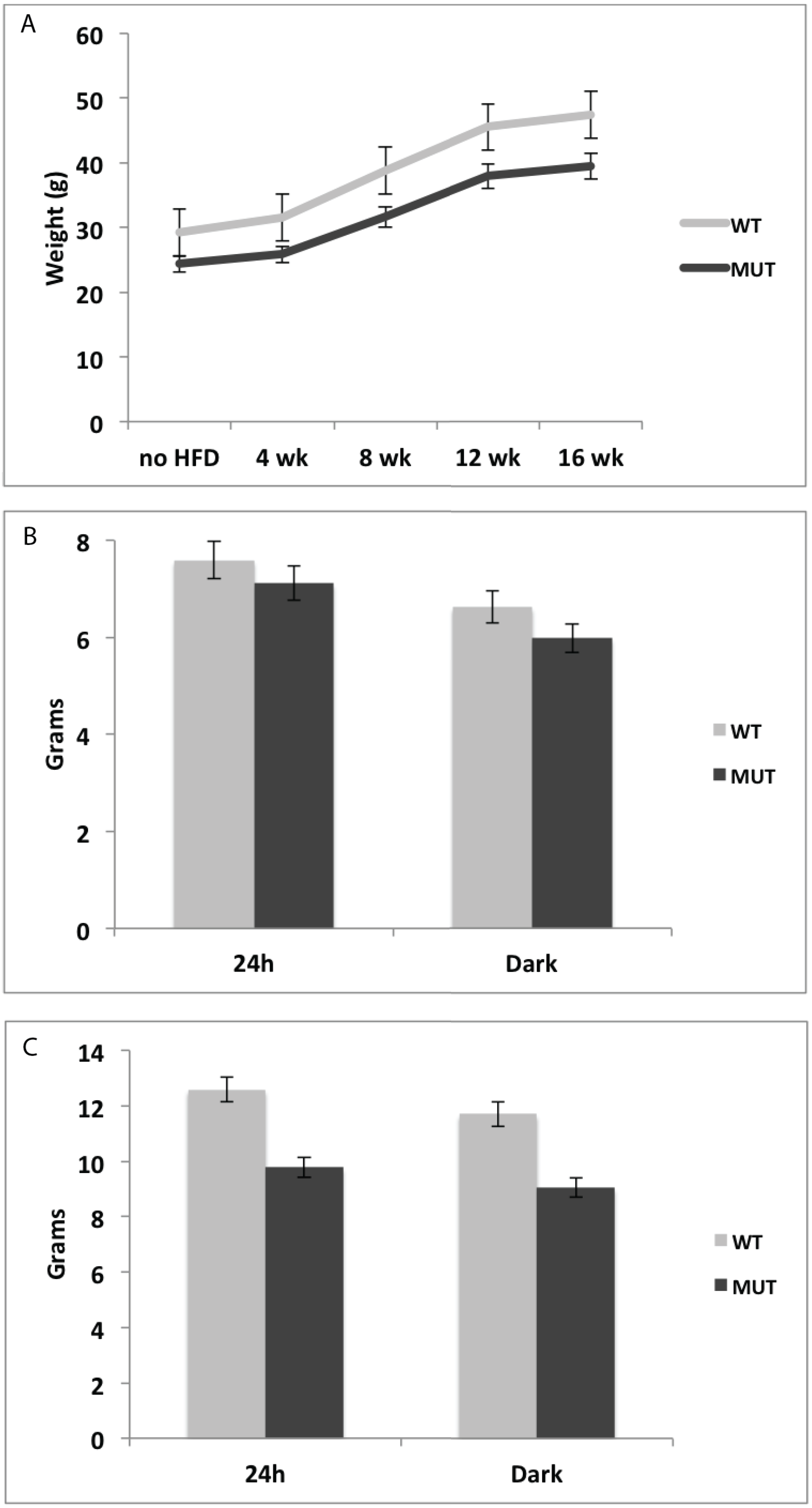


Figure 45. CBP mutants are resistant to diet-induced obesity.
CBP mutants are lighter and have less fat mass than wild type controls. A) Body weight. B) Food intake after 4 weeks on HFD. C) Food intake after 16 weeks on HFD.

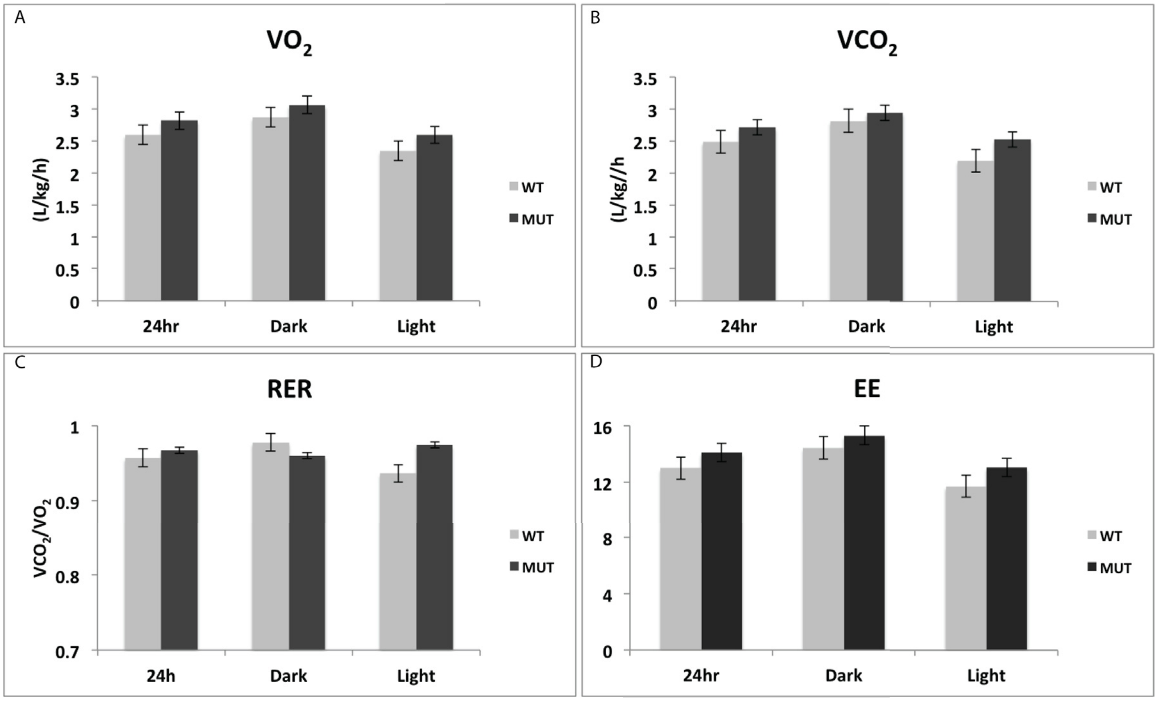


Figure 46. CBP mutants possess enhanced metabolism compared to controls. Assessment of metabolic parameters following consumption of HFD over time period of 4 weeks. A) Oxygen consumed. B) Carbon dioxide produced. C) Respiratory exchange ratio. D) Energy expenditure.

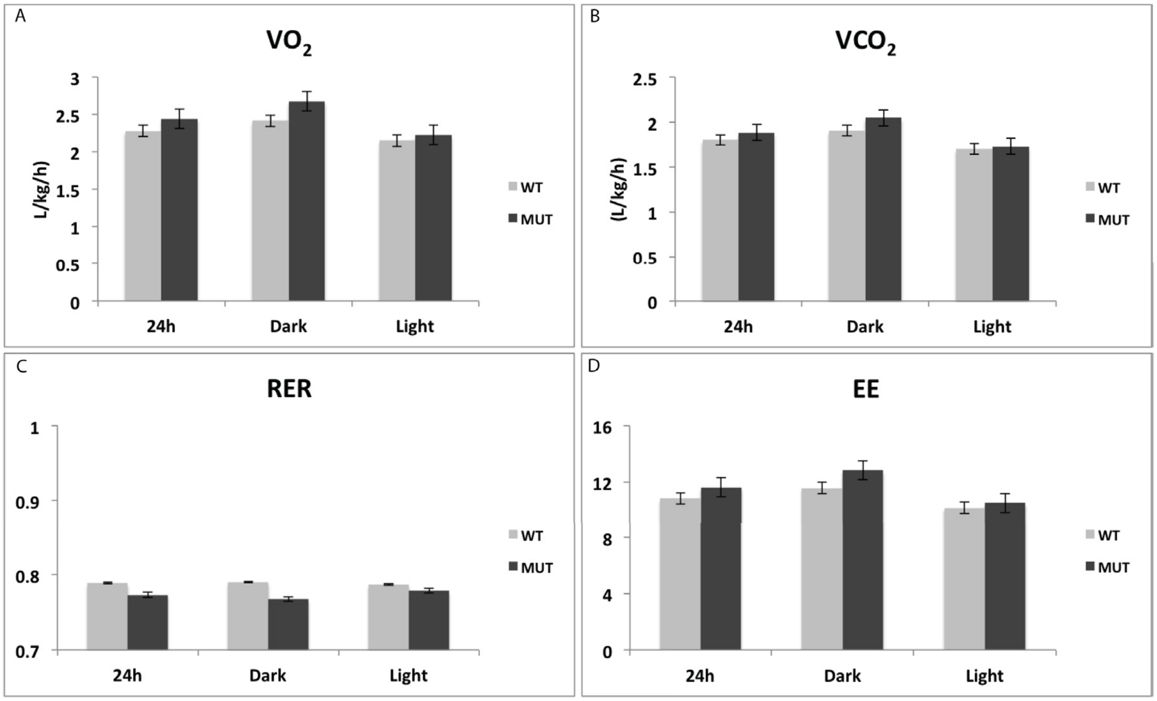


Figure 47. CBP mutants display greater energy expenditure relative to controls. Assessment of metabolic parameters following consumption of HFD over 16-week time course. A) Oxygen consumed. B) Carbon dioxide produced. C) Respiratory exchange ratio. D) Energy expenditure.

6. Appendix II: Expression, Purification, and Biophysical Characterization of CREB and CREM

6.1 Methods for expression and purification of CREB in bacterial cells

6.1.1 Cloning and expression of CREB

The full-length *Mus musculus CREB1A* gene was cloned into the TOPO vector (Invitrogen) following PCR amplification. These primers were used for PCR amplification: forward 5’-GAGAGCTCATGACCATGGAATCTGGAGCA-3’ and reverse 5’-GAGCGGCCGCTCATTAATCTGATTTGTGGCAGTAAAGG-3’. Plasmid DNA containing the cloned gene was purified and sequenced. Subsequently, both the insert and vectors pET-28a and pET-42a (Novagen) were digested with *NotI* and *SacI* (New England Biolabs). The digested ends of insert and vector were joined using T4 liagase (New England Biolabs), and following heat inactivation for 15 minutes at 60°C, ligation products were transformed into DH5α cells (New England Biolabs). Several colonies were selected for isolation of plasmid DNA using the Fermentas Miniprep Kit (ThermoScientific). DNA sequencing was used to confirm the presence of the gene in the correct reading frame. The full-length *Mus musculus CREB1A* gene was also cloned into other vectors and used with different bacterial strains and expression systems (Table 6). All methods for cloning CREB in bacterial cells were similar to those listed in this section with the exception of the use of different restriction enzymes and antibiotics based on the vector.

Plasmid DNA of CREB was used to transform *E. coli* BL21 (DE3) cells (Novagen). Cells were grown overnight at 37°C in LB medium supplemented with 100 μg/mL kanamycin and used to inoculate cultures the following morning. Eight liters of bacterial cells were grown at 37°C in LB medium supplemented with 50 μg/mL kanamycin to 0.8 at A_{600}. Protein expression was induced with the addition of 0.4 mM IPTG and growth was continued at 18°C overnight. Cells were harvested by centrifugation and washed with 50 mM sodium phosphate pH 8.0, 300 mM NaCl, 10 mM imidazole, and 10 mM β-mercaptoethanol. Following this wash step, cells were subsequently pelleted and stored at −80°C.

6.1.2 Protein purification of CREB

Cell pellets were resuspended in buffer A (50 mM sodium phosphate pH 8.0, 300 mM NaCl, 10 mM imidazole, and 10 mM β-mercaptoethanol,) and lysed by microfluidization. The lysate was fractionated into the soluble cytoplasmic fraction by centrifugation at 12,000 rpm at 4°C for 30 minutes. The soluble cytoplasmic fraction was then filtered using 0.2 μm filter. Supernatant containing CREB protein was loaded onto a HisTrap HP 5 mL column (GE Healthcare) pre-equilibrated with buffer A. CREB was eluted using an imidazole gradient with 0-100% elution buffer B (50 mM sodium phosphate pH 8.0, 300 mM NaCl, 250 mM imidazole, 10 mM β-mercaptoethanol). All fractions were collected and subjected to SDS-PAGE analysis (Figure 48). Fractions were pooled for dialysis against 1 L of dialysis buffer (20 mM Tris pH 8.5, 50 mM NaCl, and 10 mM β-mercaptoethanol). Dialysis buffer was filtered and used as binding and equilibration buffer for anion exchange chromatography. If any protein remained following dialysis, it was applied to a pre-equilibrated Source Q column (GE Healthcare).

Unbound protein was removed with 5 column volumes of binding buffer and target protein eluted with a gradient up to 100% of buffer containing 20 mM Tris pH 8.5, 1 M NaCl, and 10 mM β-mercaptoethanol over 20 column volumes (Figure 49). Fractions containing protein were pooled and concentration was attempted.

6.2 Methods for cloning, expression, and purification of CREB in insect cells

6.2.1 Production of bacmid DNA and transformation into DH10Bac cells

Full-length CREB was codon optimized for expression in insect cells (Figure 50), and cloned into the pFastBacHTC vector using *SalI* and *HindIII* as restriction sites by GeneScript. Stocks of pFastBacHTC-CREB were prepared using sterile ddH_2O. DH10Bac competent cells were thawed on ice before incubation with plasmid DNA. DNA at a concentration of 1 ng/μL was added to 100 μL of cells. Incubation proceeded for 30 minutes on ice, followed by heat shock at 42°C for 45 seconds, and cells were then transferred to ice for 5 minutes. S.O.C. medium (Invitrogen) was added (900 μL) to bring final volume to 1 mL, and cells were grown at 37°C with agitation at 220 rpm for 60 minutes. LB agar plates containing 7 μg/mL gentamycin, 50 μg/mL kanamycin, 10 μg/mL tetracycline, 100 μg/mL 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal), and 40 μg/mL IPTG were plated with either 100 μL of cells or 10 μL of cells mixed with 90 μL of S.O.C. media. Plates were incubated at 37°C for 48 hours and monitored for the appearance of blue or white colony growth. White colonies were selected and re-streaked onto fresh agar plates for 48 hours to ensure the colonies contained true transformants and to select a single colony.

6.2.2 Colony PCR

PCR reactions were prepared in thin-walled PCR tubes (Sigma) with all of the necessary components (Table 5). M13 forward 5’-GTTTTCCCAGTCACGAC-3’ and reverse 5’-CAGGAAACAGCTATGAC-3’primers were used. A 20 µL tip was used to scrape cells from agar plates and this tip was incubated for 5 minutes in PCR tube with the other reaction components before pipetting to evenly distribute cell into the solution. The reaction proceeded with an initial denaturation step at 93°C for 3 minutes to lyse cells, followed by thirty cycles of denaturation at 94°C for 30 seconds, annealing at 55°C for 30 seconds, extension at 72°C for 3 minutes and 30 seconds (60 seconds per kb length of PCR product), and final extension step at 72°C for 5 minutes. A 0.7% agarose gel (w/v) was used to visualize PCR products and confirm presence of band at expected lengths.

Table 5. Reaction conditions for colony PCR.

Reaction Component	Amount Used	Supplier
ddH_20	41.4 µL	
10X Reaction Buffer	5.0 µL	Invitrogen
10 mM dNTP Mix	1.0 µL	New England Biolabs
50 mM $MgCl_2$	1.5 µL	New England Biolabs
40 µM Forward Primer	0.3 µL	Integrated DNA Tech.
40 µM Reverse Primer	0.3 µL	Integrated DNA Tech.
Platinum Taq (5U/µL)	0.5 µL	Invitrogen

6.2.3 Isolation of Bacmid DNA

Cultures of DH10Bac cells (10 mL) were prepared with LB media supplemented with 7 µg/mL gentamycin, 50 µg/mL kanamycin, 10 µg/mL tetracycline, and inoculated for each bacmid and grown overnight at 37°C for 16 hours at 220 rpm. Cells were harvested using centrifugation at 3500 × g for 10 minutes. Qiagen miniprep kit was used to isolate bacmid DNA. Buffer P1 (250 μL) was used to resuspend the cells and transferred to microcentrifuge tube. Buffer P2 (250 μL) was added to lyse cells and inverted gently 6 times to ensure homogeneity. The suspension was incubated at RT for 5 minutes before 350 μL of buffer N3 was added and inverted gently 6 times to precipate DNA. Cellular debris was pelleted at 12,000 × g on a microcentrifuge. Pre-chilled isopropanol (700 μL) was added to clean microcentrifuge tubes before the addition of 750 μL supernatant. Tubes were inverted 6 times before incubating on ice for 10 minutes, followed by a centrifugation at 12,000 × g for 15 minutes to pellet DNA. The supernatant was removed by gently pipetting so as not to disturb pellet. Pre-chilled 70% ethanol (500 μL) was added to wash the pellet by inverting tube 6 times. Tubes were centrifuged at 12,000 × g for 5 minutes and this wash step, followed by centrifugation, was repeated once before drying DNA pellets in sterile tissue culture hood for approximately 30 minutes. DNA was dissolved in 40 μL of TE buffer and concentrations were determined using the Nanodrop. Bacmid DNA was stored at 4°C for no longer than two weeks.

6.2.4 Baculovirus production

Sf9 cells were seeded onto 6-well plates at a density of 1×10^6 cells in a total volume of 2 mL and grown at 27°C overnight. The following day plates were typically 70-80% confluent. For each transfection, liposome-DNA complexes were prepared: 6 μL of Cellfectin I (Invitrogen) was added to 100 μL of Grace's insect cell media (Invitrogen), and 3 μL of of bacmid DNA were added to 100 μL of Grace's insect cell media in sterile microcentrifuge tubes. The lipid mixture was added to the DNA mixture and incubated for 30 to 45 minutes before the addition of 800 µL of Grace's insect cell media. Grace's insect cell media was added to each well to wash the cells and media was removed from each well by aspiration. Liposome-DNA complexes (1 mL) were added to the cells and allowed to incubate for 5 hours at 27°C. Media was removed by aspiration and 2 mL complete media was added, Sf900IIISFM media (Invitrogen) supplemented with antibiotics and returned to 27°C. After 72 hours, media containing P1 virus was transferred to a microcentrifuge tube. Cells were pelleted for 5 minutes at 500 × g and 2% FBS was added to the supernatant containing P1 virus and stored at 4°C. To amplify P1 virus for P2 virus production, cells were seeded to 11×10^6 in a volume of 10 mL in a T75 flask. Cells were incubated overnight at 27°C and typically reached 70-80% confluency the following day. P1 viral stock (400 µL) was added to infect for approximately 65 hours. Media was removed and cells pelleted at 500 × g for 10 minutes. FBS was added to 2% to the supernatant containing P2 viral stock and stored at 4°C. P3 viral stock was propagated as outlined with the only exception being that P2 viral stock was used to generate P3 and only 250 µL were used to initiate infection. P3 viral stocks were used for protein expression.

6.2.5 Maintenance of insect cells

All cell maintenance was performed using aseptic technique under a sterile tissue culture hood. Cell density was determined using a hemacytometer. Sf9 insect cells were thawed from liquid nitrogen stocks and used to initiate new maintenance cultures. A 1.5 mL cryovial (Nalgene) containing cells at a density of 10×10^6 cells/mL was placed in a 15 mL conical tube and cells were pelleted for 5 minutes at 500 × g. The vial was removed with tweezers and media containing DMSO was aspirated. Cells were resuspended with 1 mL of Sf900IIISFM media and transferred to a 25 mL spinner flask. An additional 1 mL of media was added to resuspend any remaining cells and the total volume in flask was brought to 15 mL with additional media. The spinner flask was placed on a cell stirrer at a stirring rate of 85 rpm. Once cells reached a density greater than 2×10^6 cells/mL, additional media was added and growth continued until cells reached a density of approximately 3.5×10^6 cells/mL and then transferred to 125 mL spinner flask and culture passaged to a density of approximately 1×10^6 cells/mL. Subsequently, insect cells were transferred into larger spinner flasks once they reached a certain density. Typically cells were transferred into 125 mL spinner flask, followed by transfer to 500 mL spinner flask and then moved to a 1 L or 3 L spinner flask. Aeration was only used with 1 L or 3 L spinner flasks. As such, cells were generally passaged at least three to four times before protein expression was performed. Cells were discarded after reaching a passage number of 30. Following use, spinner flasks were treated with a 10% solution of bleach and washed thoroughly before sterilization. Flasks were autoclaved for a minimum of two dry cycles of 50 minutes of sterilization, followed by 15 minutes of drying time. Caps were loosened for autoclaving but tightened before removing flask from autoclave and stored until use.

6.2.6 CREB expression in insect cells

For protein expression, 2.5 L of Sf9 cells were grown to a density of 3.5×10^6 cells/mL in a 3 L spinner flask. Before protein expression was initiated, 30-40 mL of cells were removed from the flask to continue a maintenance culture. P3 viral stock in the amount of 50 mL was used for protein expression, and 40 μL of a 10 mM stock of PKA inhibitor H-89 [228] were mixed with media, and added at the same time. Cell density was monitored but typically remained constant after viral infection. Cells were harvested 62 to 67 hours post-infection by centrifugation for 15 minutes at 500 × g. Insect cell pellets were stored at −80°C until use.

6.2.7 Protein purification of CREB

Insect cell pellets were thawed in a beaker containing room temperature water and resuspended in buffer containing 50 mM sodium phosphate pH 8.0, 300 mM NaCl, 10 mM imidazole, 0.1% Triton X-100, 10 mM β-mercaptoethanol, with one dissolved Complete EDTA-free protease inhibitor tablet (Roche). Sonication was used to lyse cells at a 25% duty cycle for 1min with three second pulses and a one second pause in between pulses. Cellular debris was pelleted by centrifugation at 18,000 rpm at 4°C for 30 minutes using a Sorvall SS-34 rotor. The soluble cytoplasmic fraction was filtered with 0.2 μm SFCA-PF syringe filters (Corning) followed by another filtering step using 0.2 μm Stericup filters (Millipore). CREB protein lysate was incubated with 1 mL of a 50/50 slurry of Ni-NTA agarose beads (Qiagen) for 2 hours at 4°C with agitation. Beads were pelleted at low speed using centrifugation and washed with binding buffer (50 mM sodium phosphate pH 8.0, 300 mM NaCl, 10 mM imidazole, and 10 mM β-mercaptoethanol), for three 5 mL washes. Protein was eluted stepwise with 1 mL of buffer containing 50 mM sodium phosphate pH 8.0, 300 mM NaCl, and 10 mM β-mercaptoethanol with varying concentrations of imidazole (50 mM, 100 mM, 150 mM, or 250 mM) for a total of 12 washes. Fractions were analyzed by SDS-PAGE (Figure 51, B) and pooled for dialysis against 1 L of dialysis buffer (20 mM Tris pH 8.0, 50 mM NaCl, and 10 mM β-mercaptoethanol). Dialysis buffer was filtered and used as binding and equilibration buffer for anion exchange chromatography. Protein was applied to a pre-equilibrated Mono Q 5/50 column (GE Healthcare). Unbound protein was removed with 5 column volumes of binding buffer and target protein eluted with a gradient up to 100% of buffer containing 20 mM Tris pH 8.0, 1 M NaCl, and 10 mM β-mercaptoethanol over 20 column volumes (Figure 51, A and C). Fractions containing protein were pooled and dialyzed sequentially into 20 mM Tris pH 8.0, 100 mM NaCl, and 10 mM β-mercaptoethanol and concentrated to approximately 10 mg/mL for storage at −80°C.

6.2.8 Screening of crystallization conditions

CREB protein was screened for crystallization at a concentration of approximately 10 mg/mL. Recombinant protein was expressed in the soluble cell lysate from insect cells. Hanging drops consisting of 300 nL of protein were mixed with 300 nL of reservoir solution using the Mosquito (TTP Labtech). The total reservoir volume was 100 μL. CREB protein and CRE DNA were mixed in an equimolar ratio (Figure 52) for the screen with Natrix Screen HT (Hampton Research). However, only CREB protein was used for the screens with MPD, PEGRx 1 and 2. Trays were incubated at 20 °C.

Table 6. Expression systems utilized for production of CREB protein.

Construct	Expression System	Problem
pET-42a-CREB	T7	Co-purified with tag
pET-28a-CREB	T7	Precipitates easily
pET-28a-CREB	RosettaGami/T7	Low expression
pFastBacHTC-CREB	Bac-to-Bac	Low expression
pDNRDual-HsCREB	Gateway	Precipitates easily

6.3 Results and Discussion

Several full-length CREB constructs were used for protein expression in bacterial cells and insect cells. Many attempts were made to increase the expression of protein in different host systems. To ensure that certain post-translational modifications were made, we switched to a eukaryotic expression system. In particular the plasmid, pFastBacHTC-CREB was codon-optimized to increase expression in insect cells. Additionally the PKA inhibitor H-89 was also used in the hopes of increasing insect cell expression [228]. Although protein was present in the soluble fraction, most constructs were not well-behaved.

Overall, protein was hard to obtain in appreciable amounts for crystallization screens and would often precipitate at different purification steps, or while concentrating protein. X-ray quality crystals were not obtained. Several crystallization attempts were made with CREB but only one screen yielded crystals; however, these crystals were not large enough for x-ray diffraction. Additionally we used the XtalPred web server to predict the crystallizability of CREB [229]. A prediction is made based on a comparison of different protein features with those present in a target database [230]. CREB contains a long disordered region as well as other features that make crystallization difficult. It belongs to the class with the highest degree of difficulty and the results are summarized in Figure 53.

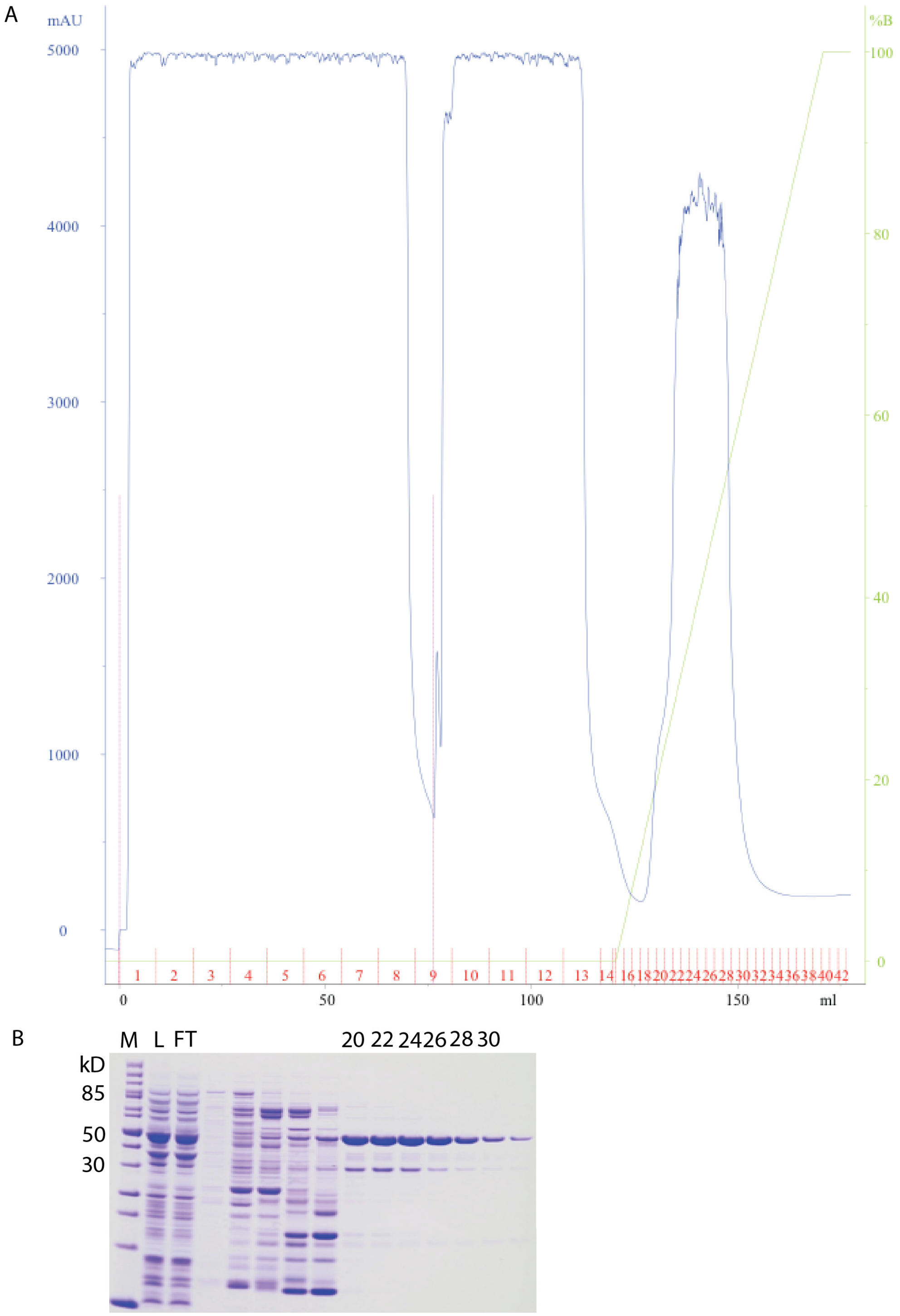


Figure 48. Ni^{2+}-affinity chromatography of CREB.

(A) Chromatogram of first step of purification. Blue line denotes absorbance at 280 nm in milliabsorbance units and green line represents gradient of buffer used for elution. Pink line denotes sample injection. (B) SDS-PAGE gel of purification stained with Coomassie blue stain. M refers to molecular weight ladder in kilodaltons and L to sample loaded onto column. Fractions corresponding to peak are shown.

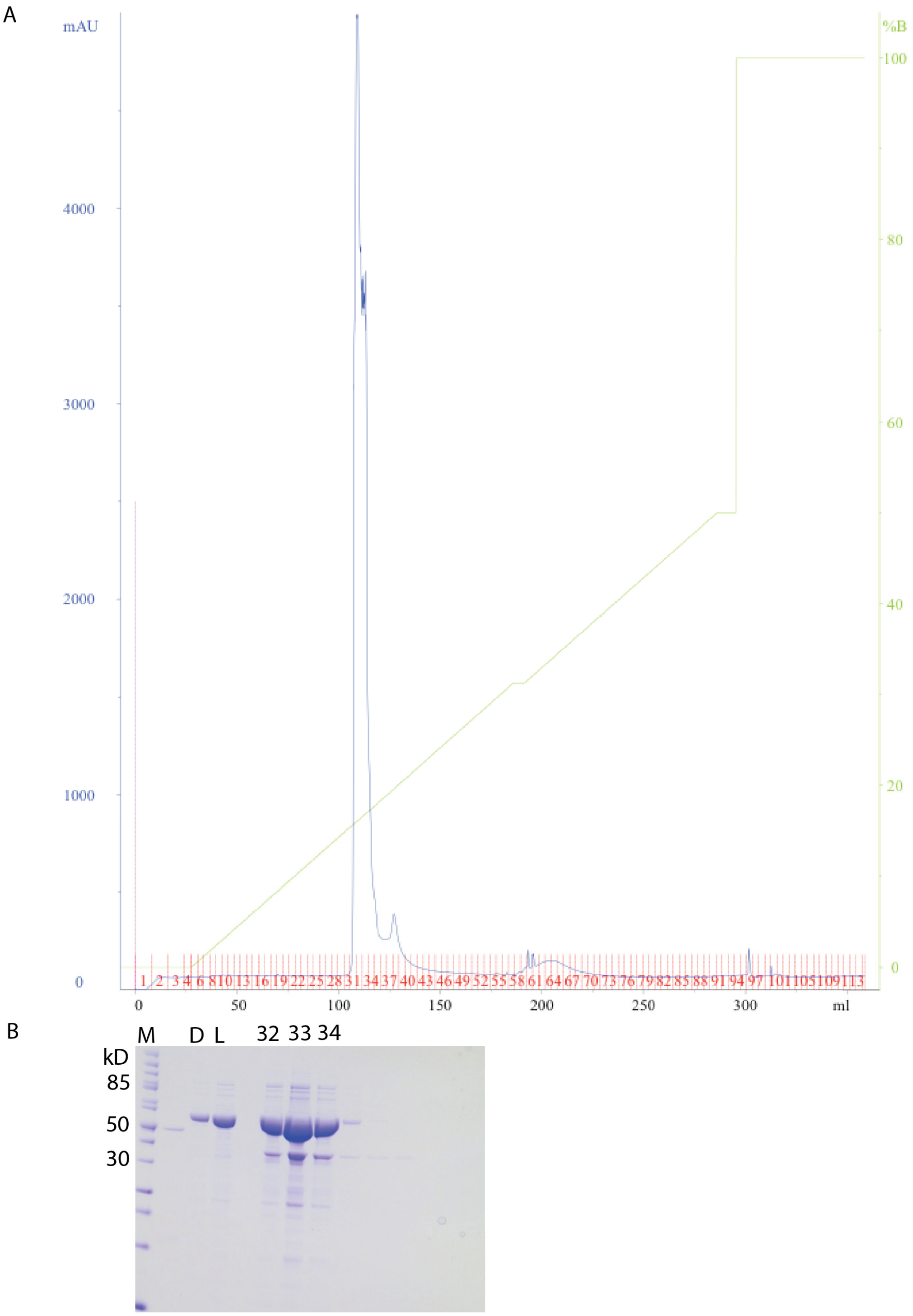


Figure 49. Anion exchange chromatography of CREB.
(A) Chromatogram of second step of purification. Blue line denotes absorbance at 280 nm and green line represents gradient of buffer used for elution with concentration listed on right axis. Pink line denotes sample injection. (B) SDS-PAGE gel of purification stained with Coomassie blue stain. M refers to molecular weight ladder, D fractions after dialysis, and L to sample loaded. Fractions corresponding to peak are shown.

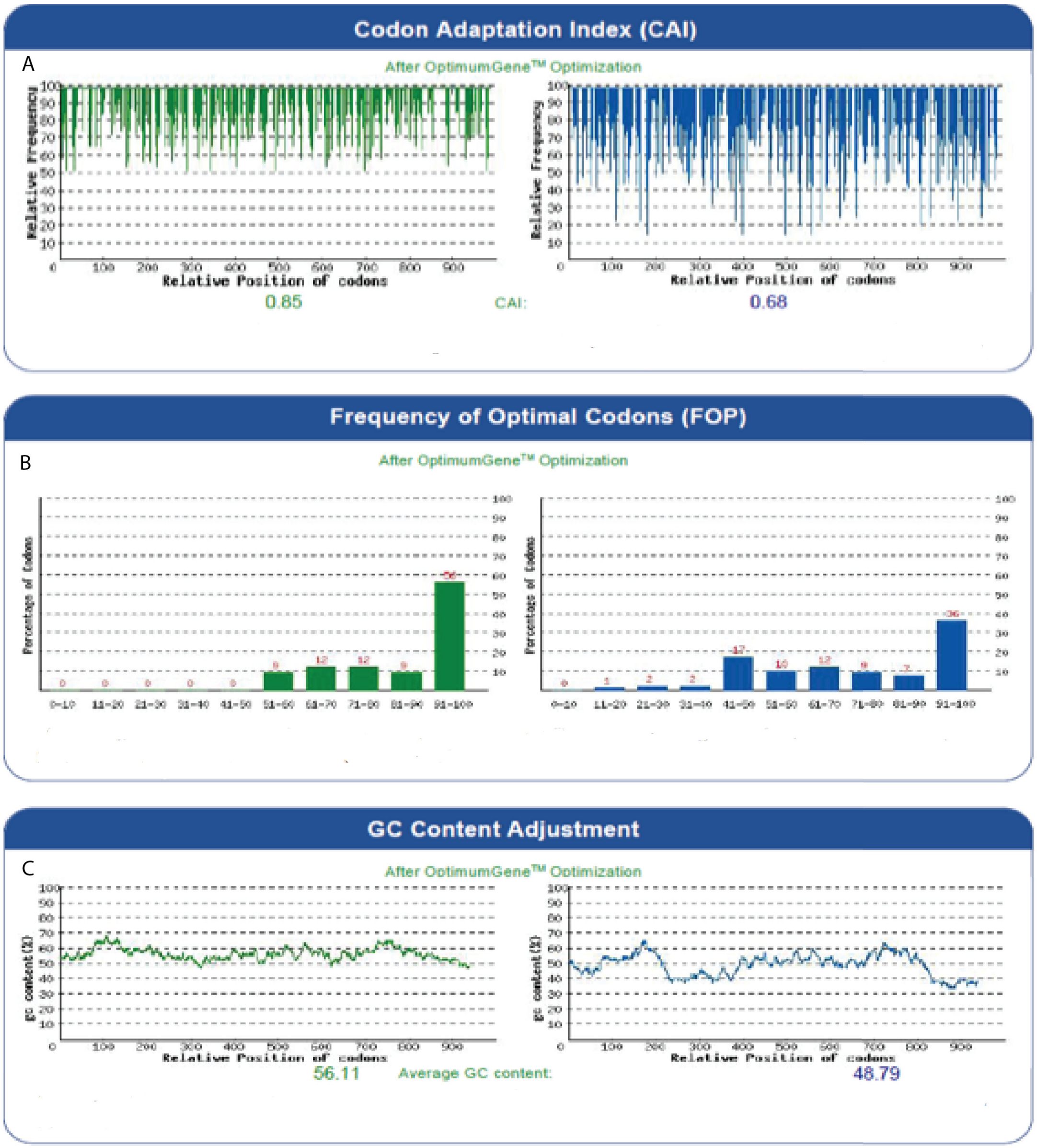


Figure 50. Codon optimization of CREB for insect cell expression.

Adjustment of codon usage bias and GC (guanine and cytosine) content. The original values are displayed in blue while values following optimization are displayed in green. (A) Distribution of codon usage frequency along length of gene. A CAI of 1.0 is considered perfect and value > 0.9 is regarded as very good n terms of gene expression level. (B) The percentage of distribution of codons in computed codon quality groups. The value of 100 is set for the codon with highest usage frequency for a given amino acid in desired organism (insect). (C) The ideal percentage range of GC content is between 30-70%. Peaks of GC content in a 60 base pair window were removed.

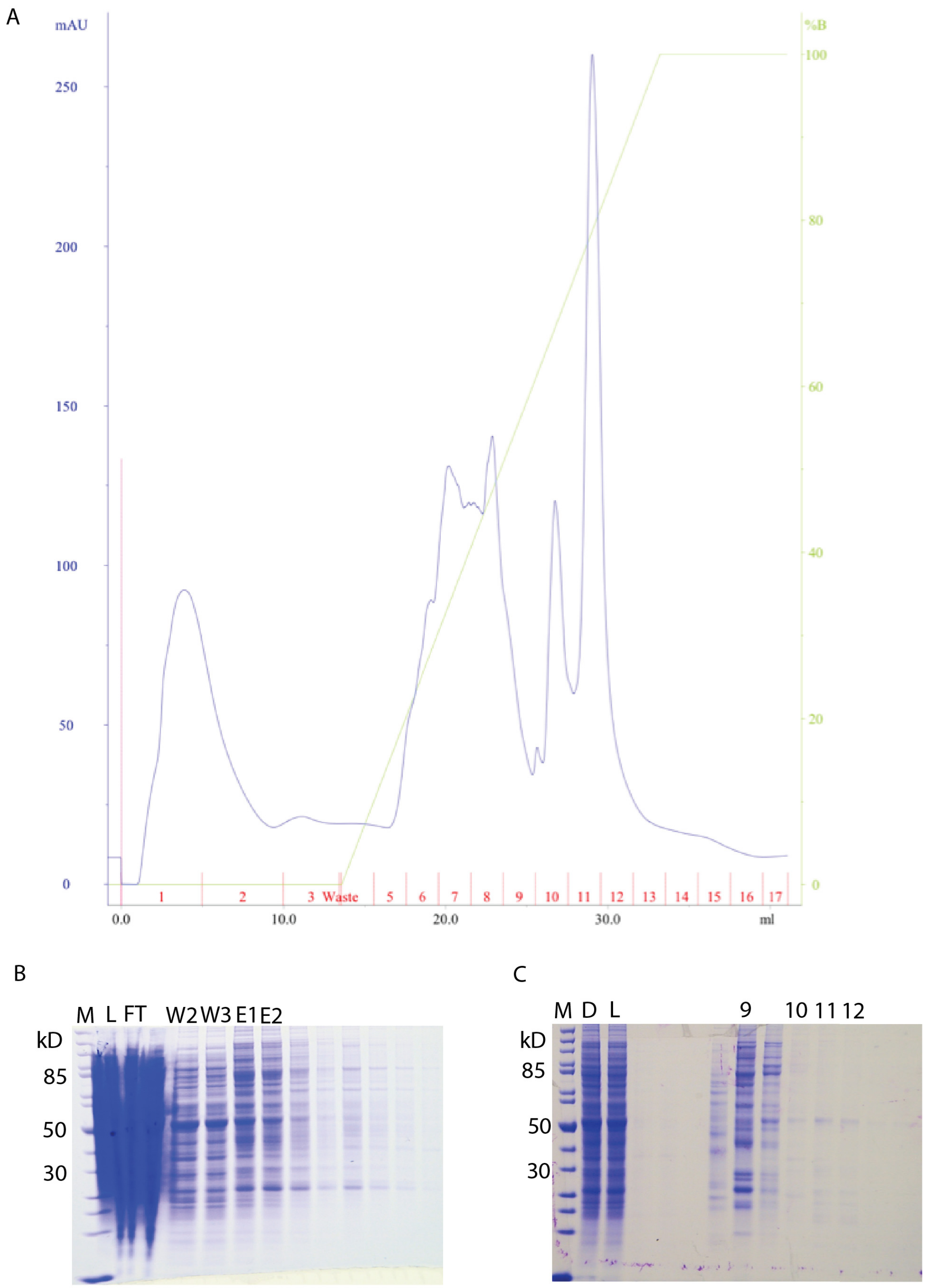


Figure 51. Purification of CREB expressed in insect cells.
(A) Chromatogram of 2^{ND} step of purification with anion exchange chromatography. Blue line denotes absorbance at 280 nm and green line represents gradient of elution buffer. Pink line denotes sample injection. (B) SDS-PAGE gel of Ni^{2+}-affinity chromatography purification. M refers to molecular weight ladder, L to sample loaded, W to wash steps, E to elution steps. (C) SDS-PAGE gel of anion exchange. D refers to pooled samples after dialysis and fractions corresponding to peaks are shown.

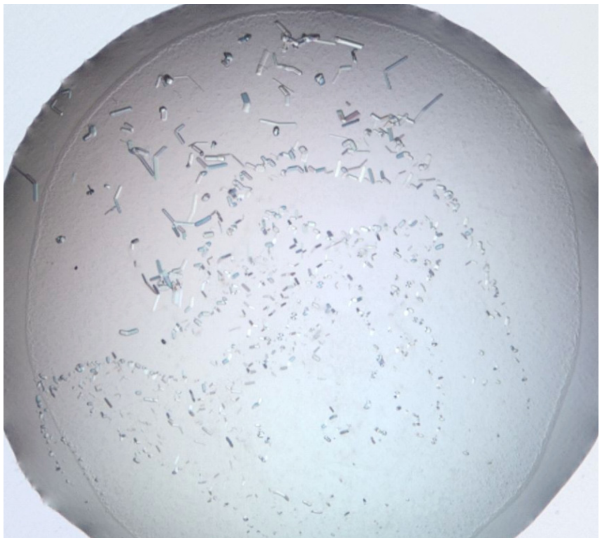
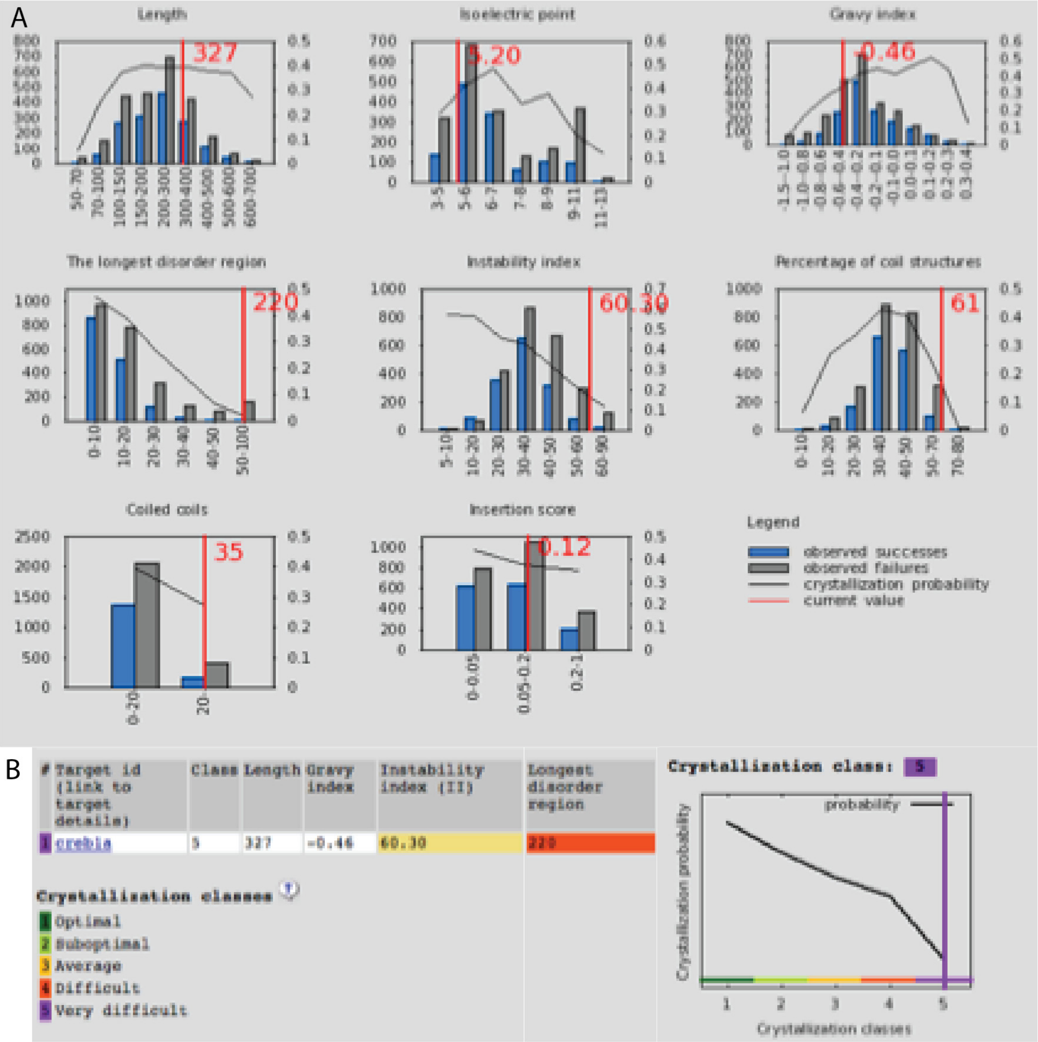


Figure 52. Crystals of CREB in complex with CRE.
CREB protein and CRE (DNA duplex) were mixed in an equimolar ratio and crystals were grown by hanging drop vapor diffusion. However, diffraction quality crystals were not obtained.



#	Target id (link to target details)	Class	Length	Gravy index	Instability index (II)	Longest disorder region
1	crebia	5	327	-0.46	60.30	220

Crystallization classes

1 Optimal
2 Suboptimal
3 Average
4 Difficult
5 Very difficult

Figure 53. Crystallization probability analysis of CREB.
(A) Histograms representing comparison of target features with distribution probabilities obtained for: length, isoelectric point, gravity index, length of disordered regions, instability, percentage of coil structures, and number of coiled coils displayed for CREB. (B) Crystallization class with total length, gravity index, instability index, and longest disordered region listed for CREB. Degree of difficulty denoted by color in legend.

6.4 Sequence alignment of CREB and CREM

The amino acid sequences of CREB and CREM were selected for sequence alignment. Protein sequences of full-length CREB1A and CREB1B, and a truncated version of CREM were aligned using Clustal Omega [162]. The figure containing the alignment was generated using the program ESPript [163].

6.5 Methods for expression, purification, and biophysical characterization of CREM

6.5.1 Expression and purification of CREM

A clone containing a truncated form of the *Homo sapiens CREM* gene (amino acid residues 21-292) was obtained from the DNASU plasmid repository. Plasmid DNA of *CREM* was used to transform *E. coli* BL21 (DE3) cells (Novagen). Cells were grown overnight at 37°C in LB medium supplemented with 100 μg/mL ampicillin and used to inoculate 6 liters of Autoinduction Media containing the following: 1% tryptone, 0.5% yeast extract, 25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 2 mM $MgSO_4$, 54 mM glycerol, 2.8 mM glucose, and 5.6 mM α-lactose. Cells were grown at 18°C, protein expression was auto-induced, and growth continued overnight. Cells were harvested by centrifugation and washed with 50 mM sodium phosphate pH 8.0, 300 mM NaCl, 10 mM imidazole, 0.1% Triton X-100, and 10 mM β-mercaptoethanol. Following this wash step, cells were subsequently pelleted and stored at −80°C.

After thawing, cells were resuspended in lysis buffer: 50 mM sodium phosphate pH 8.0, 300 mM NaCl, 10 mM imidazole, 0.1% Triton X-100, and 10 mM β-mercaptoethanol, with one dissolved Complete EDTA-free protease inhibitor tablet (Roche), and lysed using microfluidization. Cellular debris was pelleted and removed by centrifugation on a Sorvall GSA rotor for 30 minutes at 12,000 rpm. The lysate was filtered through a 0.2 μm filter and loaded onto a HisTrap HP 5 mL column (GE Healthcare) pre-equilibrated with buffer containing 50 mM sodium phosphate pH 8.0, 300 mM NaCl, 10 mM imidazole, and 10 mM β-mercaptoethanol. Unbound proteins were washed for 5 column volumes with this buffer. Target protein was eluted using a segmented imidazole gradient with 12–100% elution buffer containing 50 mM sodium phosphate pH 8.0, 300 mM NaCl, 250 mM imidazole, and 10 mM β-mercaptoethanol. First step to 12% over 20 column volumes, followed by holding step that continued at 12% for 10 column volumes, with subsequent gradient to 35% over 20 column volumes, and to 100% over 10 column volumes in final step (Figure 55, A). SDS-PAGE analysis was used to analyze fractions that contained protein (Figure 55, B). Fractions containing protein were pooled and dialyzed using 10 kD MWCO membrane (Pierce Snakeskin) against an excess of buffer containing 20 mM Tris-Cl pH 8.0, 300 mM NaCl, and 5 mM β-mercaptoethanol. Tobacco Etch Virus (TEV) protease was added at 4°C with agitation overnight. The following morning, protein was loaded onto HisTrap HP 5 mL column (GE Healthcare) pre-equilibrated with filtered dialysis buffer supplemented with 10 mM imidazole, to separate protein from tag. Elution buffer containing 20 mM Tris-Cl pH 8.0, 300 mM NaCl, 250 mM imidazole, and 5 mM β-mercaptoethanol used to elute uncleaved protein and bound tag from column. Unbound flow-through fractions containing tag-free protein were pooled and analyzed using SDS-PAGE analysis (Figure 56).

Fractions containing target protein were dialyzed twice against 2L of binding buffer (20 mM Tris-Cl pH 8.8, 50 mM NaCl, and 5 mM β-mercaptoethanol) for a total of 5 hours. The dialysis buffer was filtered through a 0.2 μm filter and used as the equilibration and binding buffer for the next step of purification. The protein was further purified using anion exchange chromatography over a Source Q column (GE Healthcare). The column was pre-equilibrated with buffer A (20 mM Tris-Cl pH 8.8, 50 mM NaCl, and 5 mM β-mercaptoethanol) before applying protein sample over the column. Unbound proteins were washed with 4 column volumes of buffer A with target protein eluted with a gradient to 100% of buffer B (20 mM Tris-Cl pH 8.8, 1 M NaCl, and 5 mM β-mercaptoethanol) over 20 column volumes (Figure 57). Fractions of protein containing ≥ 95% purity were collected and dialyzed against an excess of buffer containing 20 mM Tris-Cl pH 8.0, 300 mM NaCl, and 5 mM β-mercaptoethanol. Protein was concentrated to 1.5 mL and subjected to gel filtration chromatography using a Superose 12 10/300 GL column (GE Healthcare). Isocratic elution of protein was obtained with buffer containing 20 mM Tris-Cl pH 8.0, 150 mM NaCl, and 10 mM β-mercaptoethanol, and protein was eluted in one column volume (Figure 58). Fractions of protein corresponding to the peak were concentrated to 10-15 mg/mL, and stored at −80°C.

6.5.2 Expression and purification of TEV protease

A plasmid containing TEV protease was re-streaked onto an LB agar plate from a glycerol stock. A single colony was selected from plate and used to grow an overnight culture in Terrific Broth supplemented with 100 μg/mL ampicillin at 37°C. The overnight culture was used to inoculate 4 L of Terrific Broth and grown at 37°C until 1.0 at A_{600}. Protein expression was induced with the addition of 1 mM IPTG and growth was continued at 17°C overnight. Cells were harvested by centrifugation and washed with 50 mM sodium phosphate pH 7.5, 400 mM NaCl, 10 mM imidazole, 10% glycerol, and 5 mM β-mercaptoethanol. Following this wash step, cells were subsequently pelleted and stored at −80°C.

Bacterial cell pellets were thawed and resuspended in buffer containing 50 mM sodium phosphate pH 7.5, 400 mM NaCl, 10 mM imidazole, 10% glycerol, and 5 mM β-mercaptoethanol, and lysed by microfluidization. The lysate was fractionated into the soluble cytoplasmic fraction by centrifugation at 12,000 rpm at 4°C for 30 minutes. The soluble cytoplasmic fraction was then filtered using 0.2 μm filter. Supernatant containing Protein was loaded onto a HisTrap HP 5 mL column (GE Healthcare) pre-equilibrated with the same buffer listed above. TEV was eluted using an imidazole gradient with 0-100% elution buffer with 50 mM sodium phosphate pH 8.0, 400 mM NaCl, 250 mM imidazole, 10% glycerol and 5 mM β-mercaptoethanol. All fractions were collected and subjected to SDS-PAGE analysis (Figure 59). Fractions containing target protein were pooled and dialyzed against 2 L of 50 mM Tris-Cl pH 7.5, 50 mM NaCl, 10% glycerol, and 5 mM β-mercaptoethanol for one hour at 4°C. Dialysis continued for an additional 3 hours with fresh buffer and sample was placed at 4°C overnight.

Dialyzed sample was centrifuged at 10,000 rpm for 20 minutes using a Sorvall GSA rotor and the supernatant filtered with a 0.2 μm filter. The protein was further purified using cation exchange chromatography over a Source S column (GE Healthcare) pre-equilibrated with buffer A (50 mM Tris-Cl pH 7.5, 50 mM NaCl, 10% glycerol, and 5 mM β-mercaptoethanol), before applying protein sample over the column. Target protein was eluted with a gradient to 100% of buffer B (50 mM Tris-Cl pH 7.5, 1 M NaCl, 10% glycerol, and 5 mM β-mercaptoethanol) over 20 column volumes (Figure 60). Elution fractions corresponding to the peak were pooled and protein concentration was determined. TEV was diluted to 1mg/mL with storage buffer containing 50 mM Tris-Cl pH 7.4, 200 mM NaCl, 40% glycerol, and 5 mM β-mercaptoethanol, and stored in 1.2 mL aliquots at −80°C.

6.5.3 Screening of crystallization conditions

Recombinant protein was expressed in the soluble cell lysate from bacterial cells. CREM was screened for crystallization at concentrations between 10-15 mg/mL. Hanging drops consisting of 300 nL of protein were mixed with 300 nL of reservoir solution using the Mosquito (TTP Labtech). The total reservoir volume was 100 μL. The following screens were used: Natrix Screen HT (Hampton Research), MPD (Qiagen), Crystal Screen I and II, Salt Rx HT, and PEGRx HT (Hampton Research). CREM protein and CRE DNA were mixed in an equimolar ratio for the screens with Natrix HT, MPD, and Crystal Screen I and II, but only CREM protein was used for the screens with Salt Rx HT, PEGRx HT. Trays were incubated at 20 °C.

6.5.4 Overview and methods used for small angle X-ray scattering

Small-angle X-ray scattering (SAXS) is a biophysical technique that can provide information about the size and shape of macromolecules in solution [231-233]. A pure protein sample that is homogenous in solution is exposed to a collimated beam of X-rays, and the scattering intensity is recorded as a function of the scattering angle. In contrast to X-ray crystallography, no crystals are needed for data collection and the solution profiles are continuous and radially symmetric. One disadvantage of this method relative to X-ray crystallography is that it does not provide information about atomic coordinates. In order to account for the different orientations of macromolecules in solution, the data generated from SAXS is spatially averaged and is of low resolution [232, 234, 235]. Modeling programs are used to calculate the molecular envelope of the macromolecule *ab initio*.

Prior to SAXS, protein samples should be subjected to gel filtration to ensure that the sample is monodispersed and of high purity ($\geq$ 95%) [236]. Purity should be confirmed by SDS-PAGE analysis. It is also imperative that matched solvent blanks are used for scattering data analysis. Macromolecules in solution as well as buffer molecules will give rise to elastic X-ray scattering when bombarded. Scattering data should be obtained for both buffered protein samples as well as matched solvent in the absence of protein. Solvent subtraction is performed to remove scattering data generated solely from the solvent and obtain scattering from the macromolecule. A scattering profile is generated for the macromolecule or complex and this profile is subsequently converted into a one dimensional scattering curve I(q). The resulting scattering curve is best described as a spherically averaged molecular envelope. The pair distribution function P(r) can be calculated by indirect Fourier transform of I(q), and describes the distance between electron pairs within the macromolecule [237]. It is a space representation of the scattering data. A P(r) plot utilizing the Guinier approximation can be used to calculate the radius of gyration (R_G) and the maximum particle dimension (D_{max}) of the protein [232, 237]. The R_G refers to the distance of an object's parts from its center of gravity or from an axis.

Wide-angle X-ray scattering (WAXS) and SAXS data were collected at the National Synchrotron Light Source Beamline X9 of the BNL. Protein samples were aliquoted into PCR tubes and centrifuged at 12,000 × g on a microcentrifuge and an automated system was used for sample injection. Protein concentrations of 3 or 4 mg/mL were used. Replicates for each sample were averaged and buffer scattering profiles were used to extract buffer from the data. The ATSAS software suite was used for data analysis. The program PRIMUS was used to process data and establish parameters and cut-off points for data analysis [238]. The program GNOM was used to analyze the P(r) pair distributions as a function of D_{max} [237]. The output from GNOM was used in the *ab initio* modeling program DAMMIF to generate a molecular envelope through multiple iterations [239]. In the initial model, dummy beads act as placeholders for solvent or particles. DAMMIF assumes the protein is spherical with a radius equal to D_{max} and the interior is filled with dummy beads. Restrictions are imposed in the solution space to ensure particle beads are connected and tightly packed. The model is optimized through simulated annealing and the process repeated until the scattering data of the modeled shape fits the experimental data. For each sample, DAMMIF generated 10 models and the program DAMAVER was used to align and average the models, to converge on the most likely conformation based on the data [240].

6.6 Results and Discussion

Little is known about CREM at the atomic level, and we had hoped it would be amenable to biophysical characterization. Crystallization attempts were made with CREM but all attempts were unsuccessful. Additionally, given its high sequence similarity to CREB (Figure 54), we hoped to assess the protein-protein interactions between CREM and CBP and apply this information to better our understanding of binding interactions between members of the bZIP family of transcription factors and coactivators such as CBP. However, using X-ray scattering data (Figure 61, A) we determined the P(r) function derived from the SAXS profile (Figure 61, B). We generated *ab initio* models of CRE (Figure 62) and CREM in complex with CRE (Figure 63).

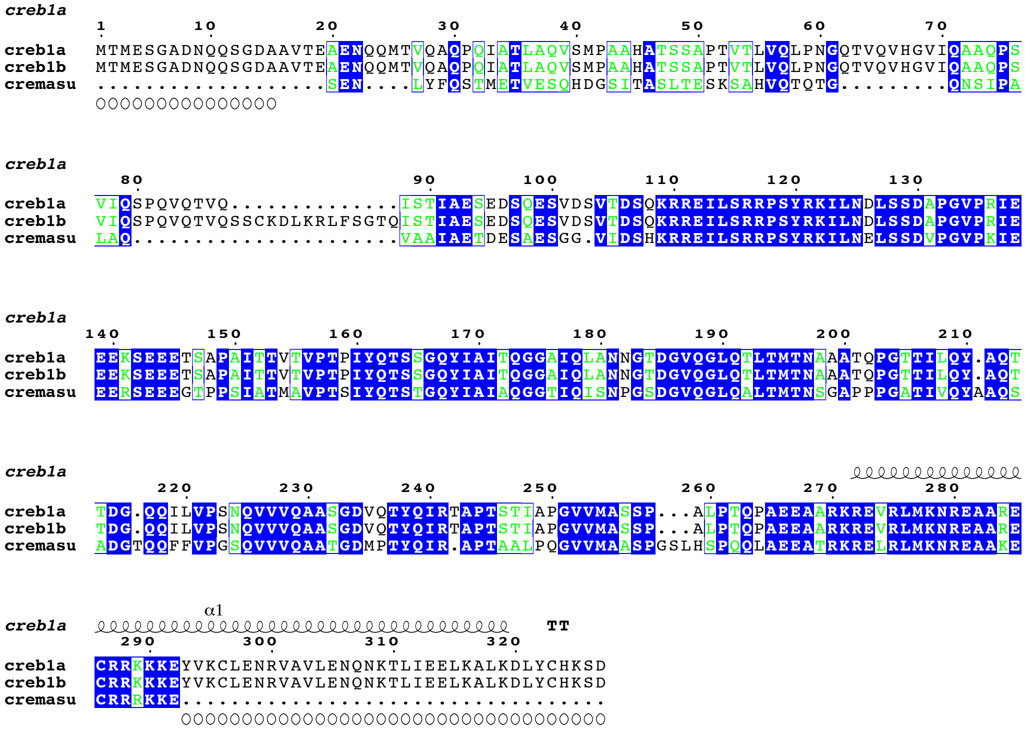


Figure 54. CREB and CREM are similar.
CREB and CREM function as activators or repressors of transcription. They bind CRE sites located within gene promoters with high. Conserved residues are displayed in blue and open circles denote deleted portions of the CREM construct used in our studies. Secondary structural elements for bZIP domain of CREB1A are displayed above the alignment.

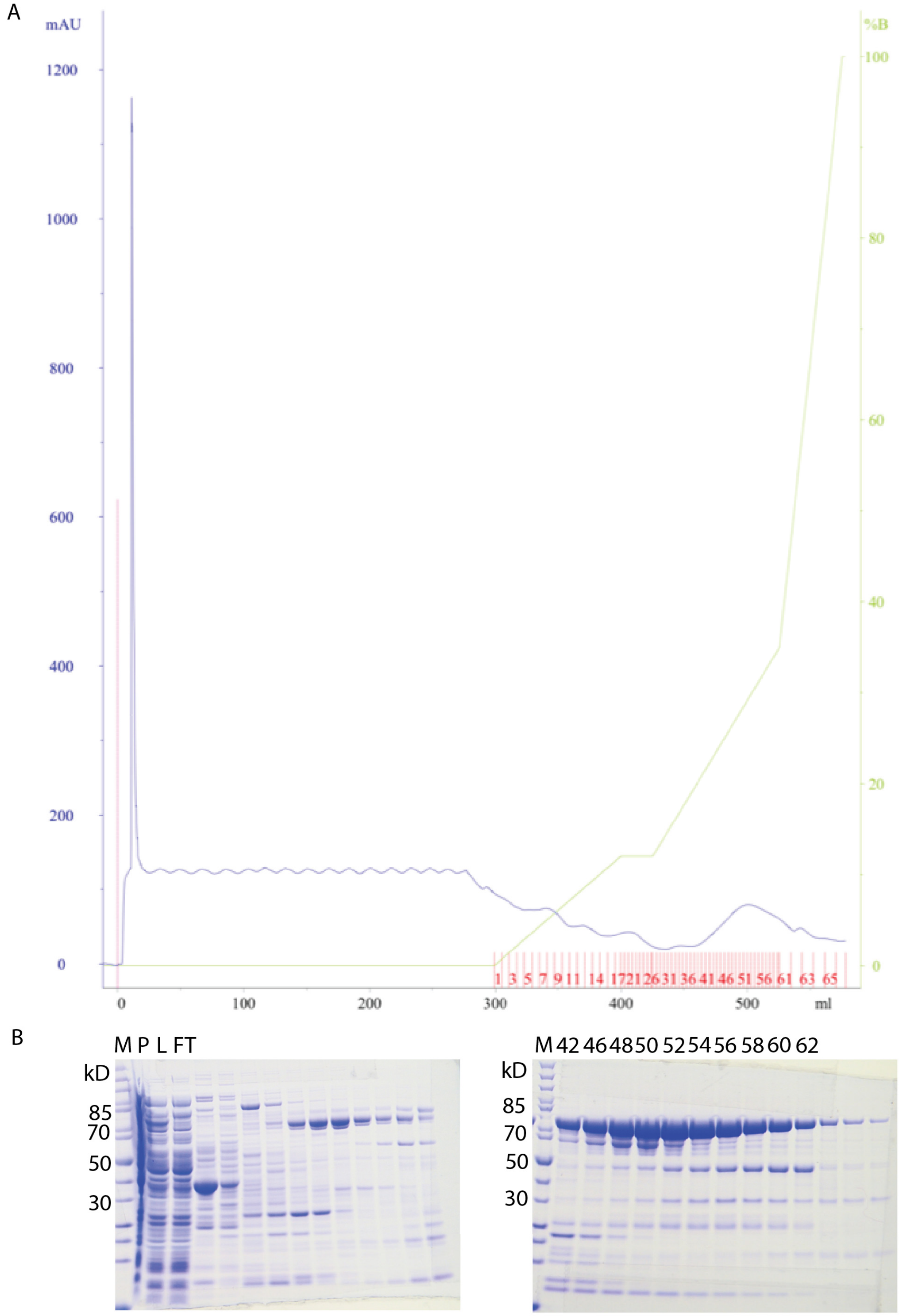


Figure 55. Ni^{2+}-affinity chromatography of CREM.
(A) Chromatogram of CREM purification. Blue line denotes absorbance at 280 nm in milliabsorbance units and green line represents concentration gradient of buffer used for elution. Pink line denotes sample injection. (B) SDS-PAGE gel of purification stained with Coomassie blue stain. M refers to molecular weight ladder, P to pellet, L to sample loaded, FT to unbound protein. Fractions corresponding to peak are shown.

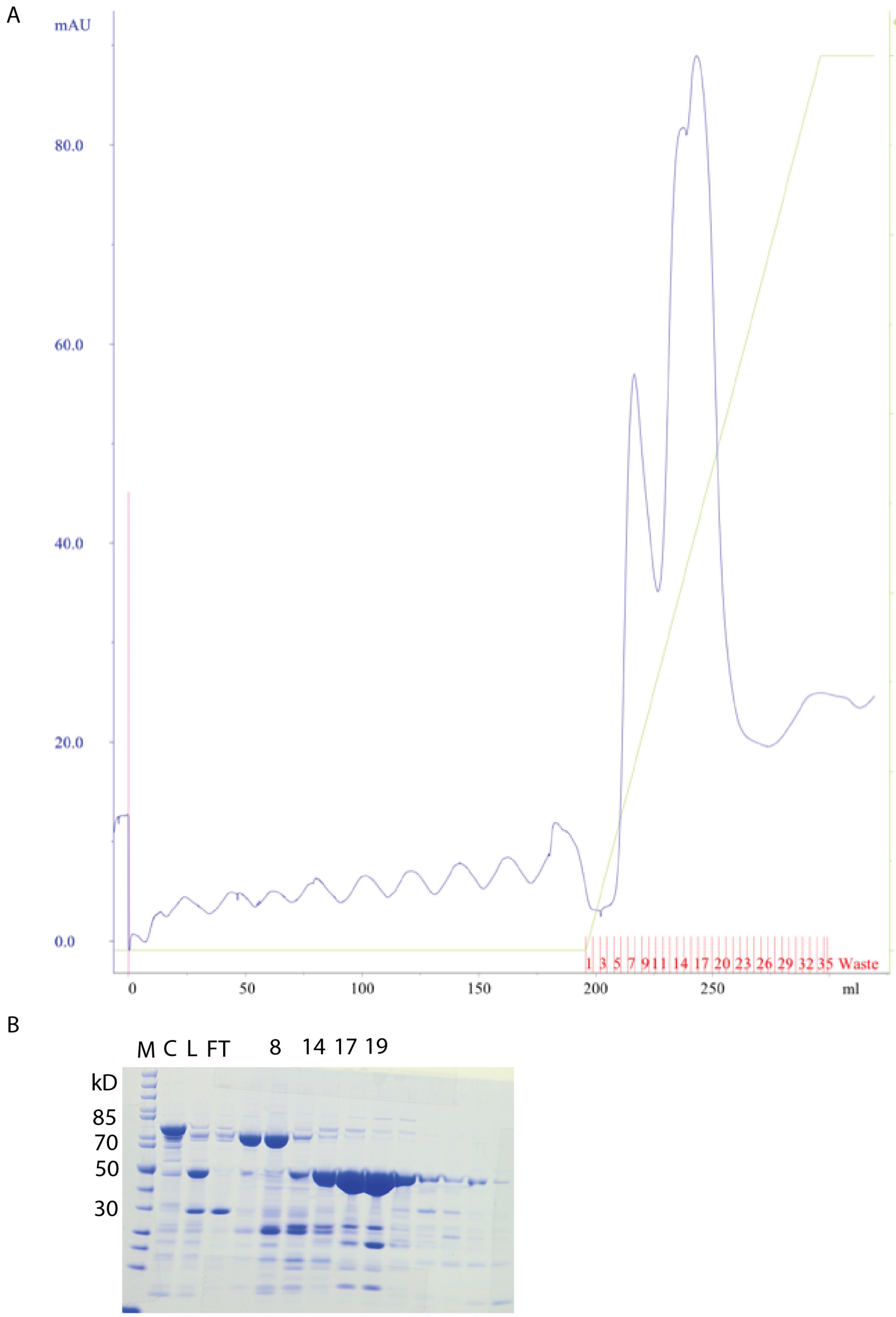


Figure 56. Ni^{2+}-affinity chromatography of CREM after TEV cleavage.
(A) Chromatogram of CREM purification. Blue line denotes absorbance at 280 nm and green line denotes concentration gradient used for elution. Pink line denotes sample injection. (B) SDS-PAGE gel of purification stained with Coomassie blue stain. M refers to molecular weight ladder in kilodaltons, C to sample before TEV protease cleavage, L to sample loaded, FT to unbound/tag-free protein. Peak fractions are displayed.

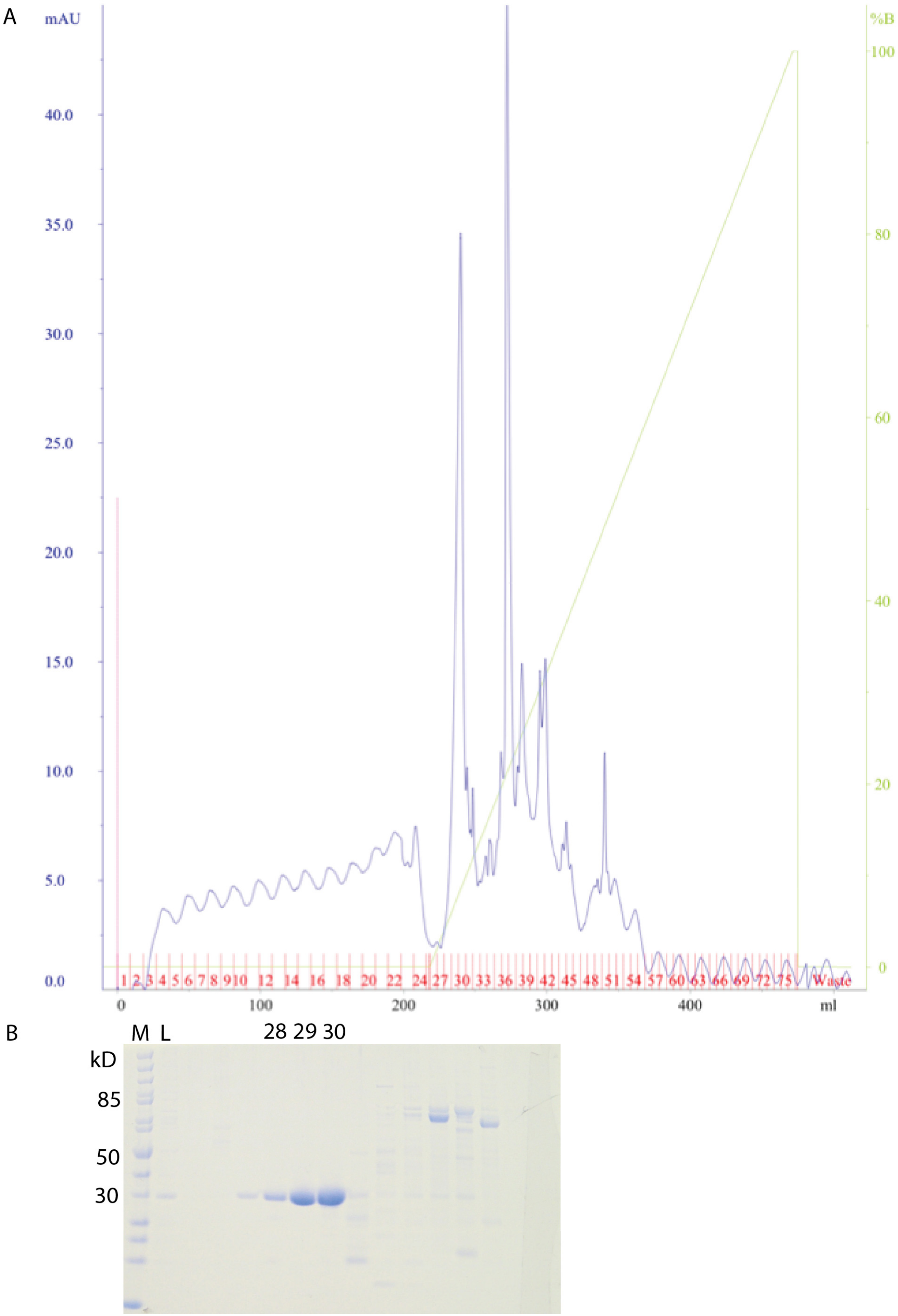


Figure 57. Anion exchange chromatography of CREM.
(A) Chromatogram of purification. Blue line denotes absorbance at 280 nm in milliabsorbance units and green line represents gradient of buffer used for elution with concentration listed on right axis. Pink line denotes sample injection. (B) SDS-PAGE gel of purification stained with Coomassie blue stain. MWT refers to molecular weight ladder in kilodaltons and L to sample loaded onto column. Fractions corresponding to peak are shown.

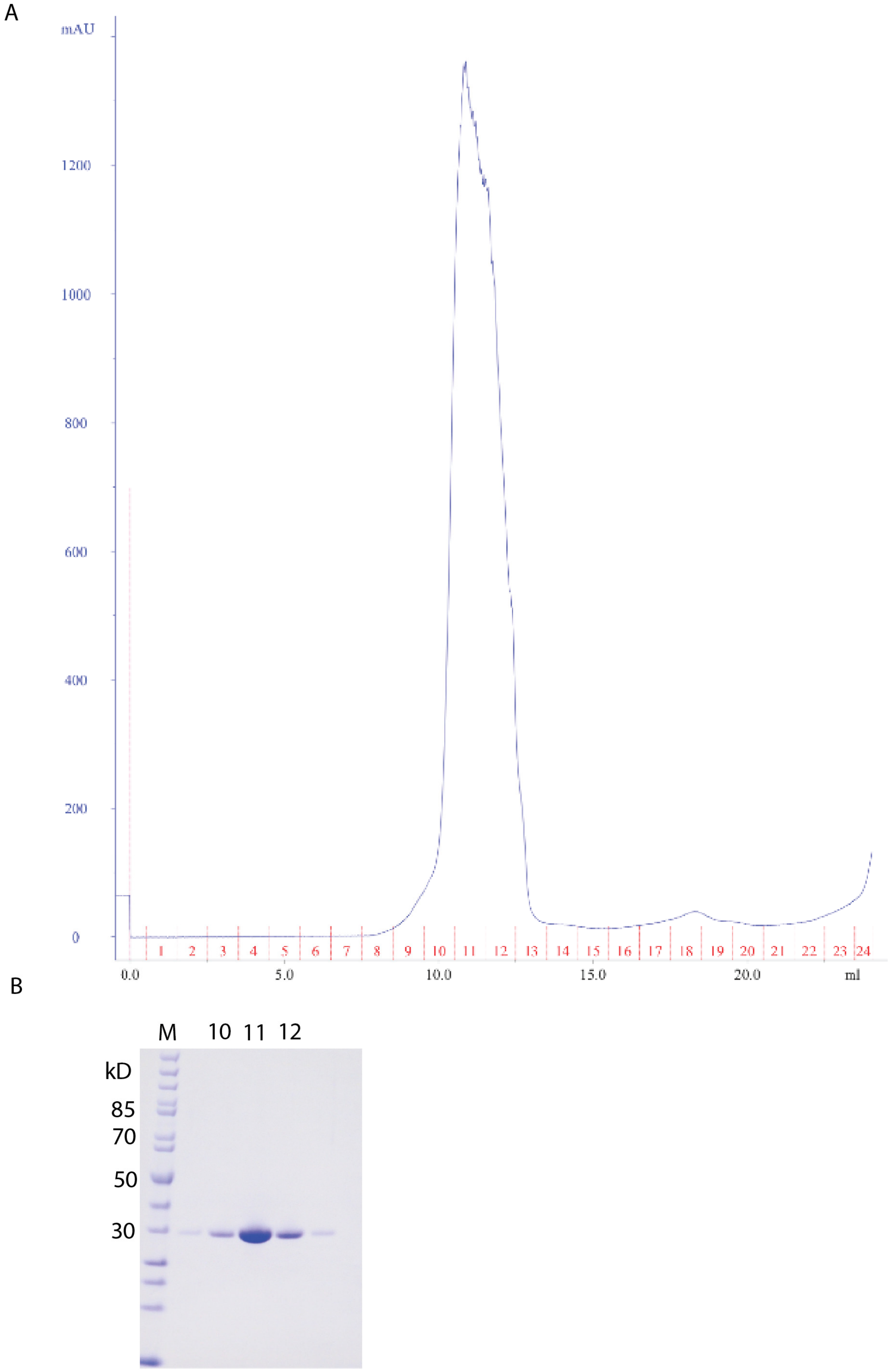


Figure 58. Gel filtration chromatography of CREM.
(A) Chromatogram of purification. Blue line denotes absorbance at 280 nm in milliabsorbance units. Pink line denotes sample injection. (B) SDS-PAGE gel of purification stained with Coomassie blue stain. M refers to molecular weight ladder in kilodaltons and fractions corresponding to peak are shown.

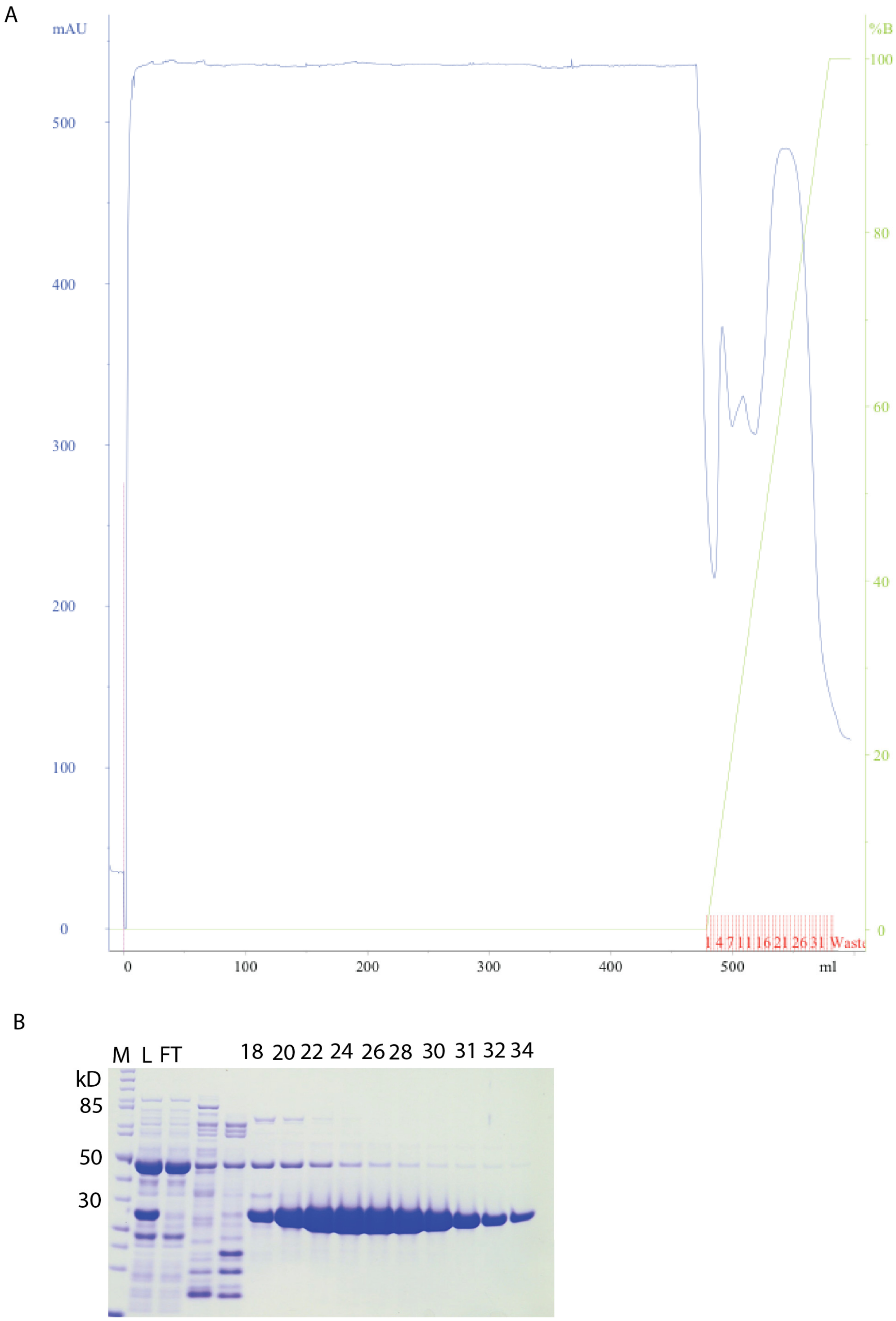


Figure 59. Ni^{2+}-affinity chromatography of TEV protease.
(A) Chromatogram displaying 1ST step of purification. Green line represents gradient of elution buffer and blue line denotes absorbance at 280 nm. Pink line denotes sample injection. (B) SDS-PAGE gel of purification stained with Coomassie blue stain. M refers to molecular weight ladder in kilodaltons, L to sample loaded, and FT to unbound protein. Fractions corresponding to peak are shown.

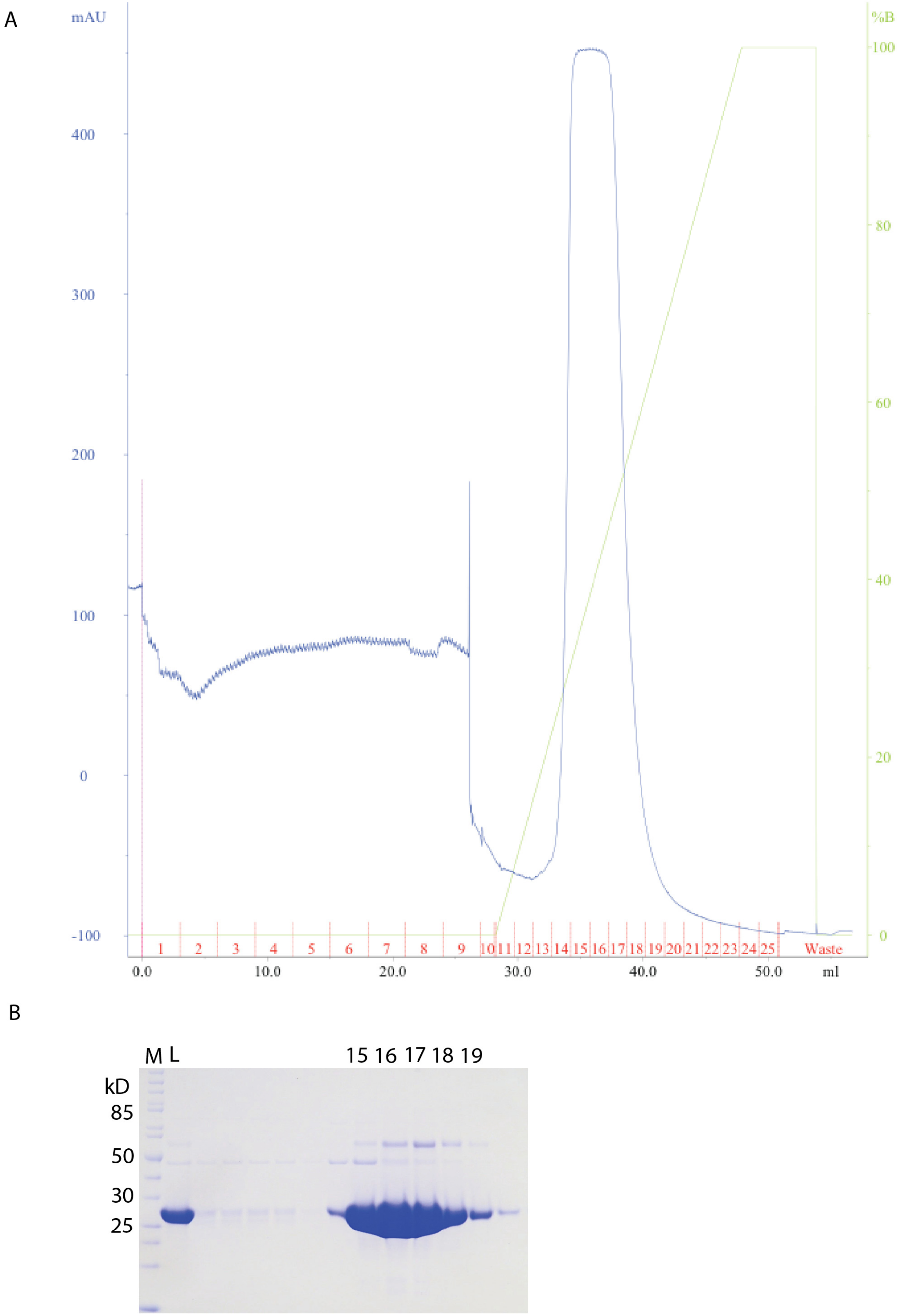


Figure 60. Cation exchange chromatography of TEV protease.
(A) Chromatogram of purification. The blue line denotes absorbance at 280 nm and green line represents gradient of buffer used for elution with concentration listed on right axis. Pink line denotes sample injection. (B) SDS-PAGE gel of purification stained with Coomassie blue stain. M refers to molecular weight ladder in kilodaltons and L to sample loaded onto column. Fractions corresponding to peak are shown.

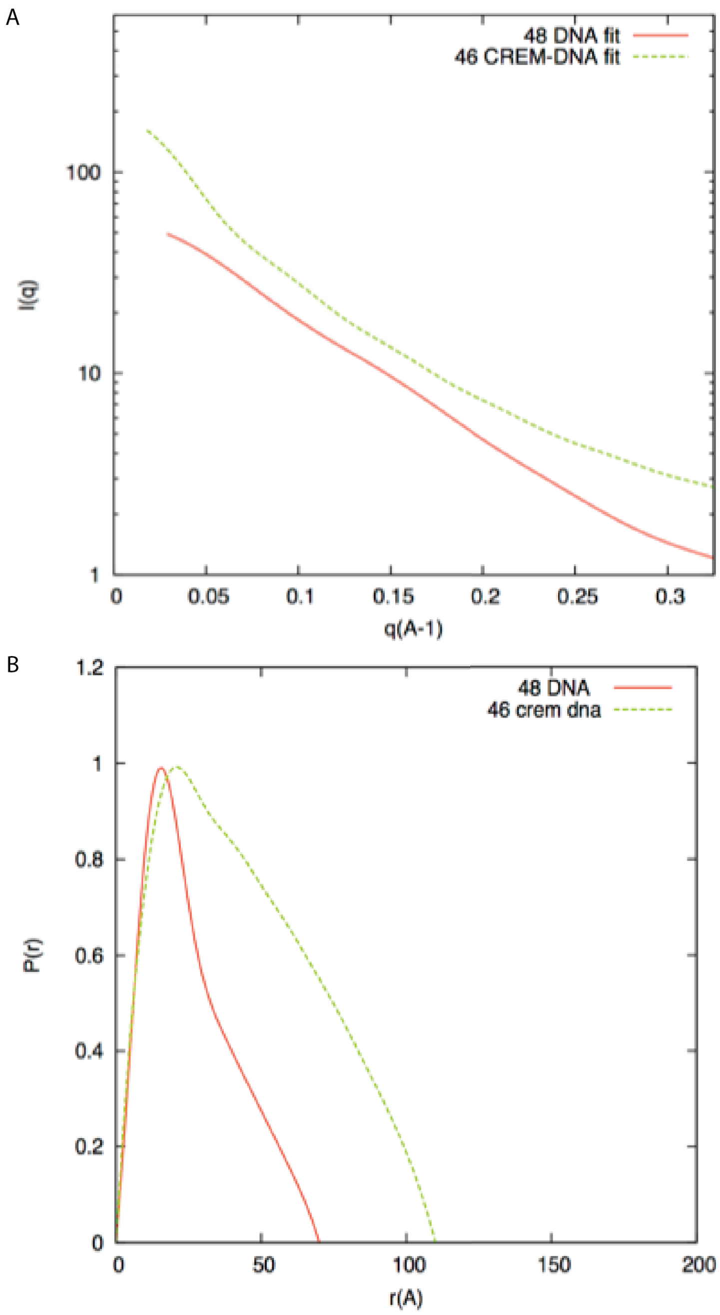


Figure 61. Scattering data and function derived from SAXS profile.
(A) Scattering intensities for CRE shown in red and CREM in complex with CRE are shown in green. (B) P(r) function derived from SAXS profile of CRE shown in red and for CREM in complex with CRE shown in green.

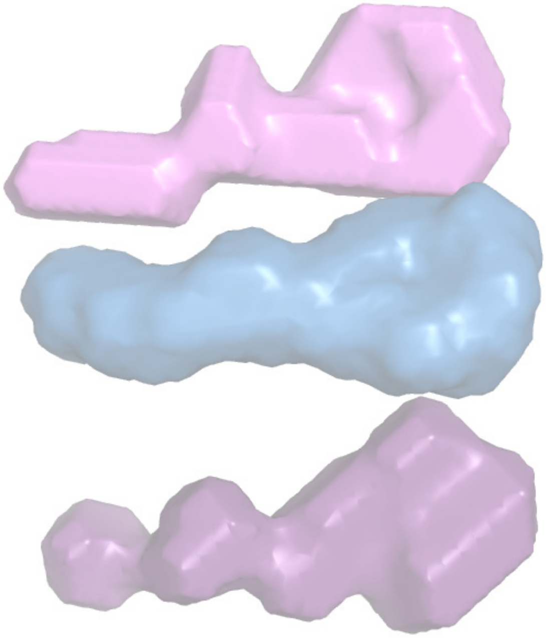


Figure 62. SAXS *ab initio* modeling results for CRE.
The *ab initio* modeling results for CRE are shown in pink and purple. The averaged *ab initio* envelope generated from the modeling results is shown in blue.

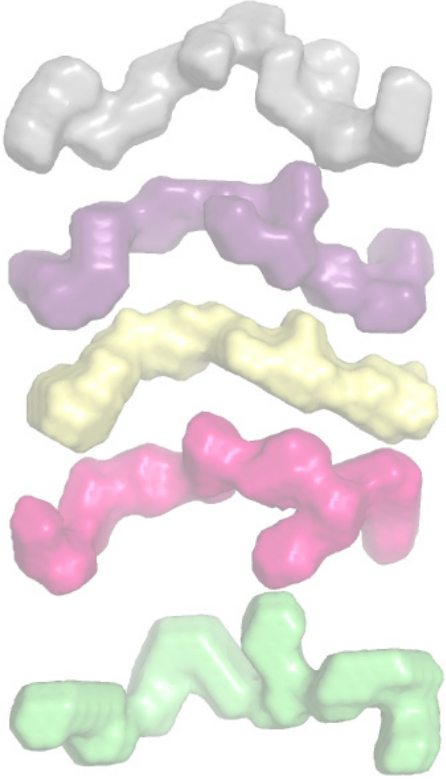


Figure 63. SAXS *ab initio* model of CREM and CRE.
The *ab initio* modeling results for CREM plus CRE are shown in gray, purple, pink and green. The averaged *ab initio* envelope generated from the modeling results for CREM in complex with CRE is shown in yellow.

7. Appendix III: Expression, Purification, and Characterization of CBP

7.1 Rationale for biophysical characterization

Although CREB and CBP play key roles in the transcriptional regulation of many different target genes, little is known about the protein interactions that drive the assembly of the CREB co-activator complex. It remains unclear which member of the co-activator complex is the predominant regulator of hepatic gluconeogenesis via CREB. Additionally, there is still much that we do not know about the binding interactions of these proteins at the atomic level.

Using small fragments of approximately 100 residues of CREB and CBP, the structures of the KID domain of CREB and the KIX domain of CBP were previously determined using NMR methods [241]. Additionally, the structure of the bZIP domain of CREB bound to a cAMP response element was resolved using x-ray crystallography [187]. While these structural determination studies provided useful information about some of the protein folds they did not provide a comprehensive picture of the long-range binding interactions that occur between these two proteins. This work provided useful insight into the mechanism of phospho-serine recognition and CREB dimerization but failed to describe an actual mechanism for the assembly of the CREB co-activator complex. We intend to use larger regions of both proteins in order to identify local protein interactions as well as determine the long-range effects of protein interactions. Binary complexes of CREB:CBP or CREM:CBP will reveal the necessary binding interactions that must occur for complex assembly and the induction of CREB transactivation. These structural studies will allow us to dissect the role of each partner in the assembly of the complex and assist us in determining which regions interact to mediate CREB-dependent gene transcription. Structural and thermodynamic characterization of the protein interactions will be carried out using biophysical methods.

7.2 Methods

7.2.1 Cloning, expression, and purification of CBP in bacterial cells

Two constructs of CBP were cloned into pET vectors (Novagen) containing truncated forms of the gene. The shorter of the two constructs corresponded to amino acid residues 302-680, while the longer of the two constructs spanned amino acid residues 1-680. PCR amplification was performed using Faststart HiFidelity Enzyme Blend (Roche) and the following primers for the short construct forward 5'-GAGAGCTC ATGGTCAATAGTTTACCTGCTT-3'and reverse 5'-GAGCGGCCGCTCAGTTACCC AGGATGCCTTGCTTAT-3'; primers for the longer construct forward 5'-GAGAG CTCATGGCCGAGAACTTGCTGGA-3' and reverse 5'-GAGCGGCCGCTCAGTTAC CAGGATGCCTTGCTTAT-3'. The PCR amplified insert was cloned into the TOPO vector (Invitrogen), and plasmid DNA containing the cloned gene was purified and sequenced. The gene and vectors were digested with restriction enzymes *NotI* and *SacI* (New England Biolabs) and cloned into pET-28a and pET-42a using these restriction sites. The digested ends of insert and vector were joined using T4 ligase (New England Biolabs). Following heat inactivation for 15 minutes at 60°C, ligation products were transformed into DH5α cells. Several colonies were selected for isolation of plasmid DNA using QiaPrep Spin Miniprep Kit (Qiagen). DNA sequencing was used to confirm the presence of the gene in the correct reading frame.

Additionally, a truncated form of the gene corresponding to amino acid residues 1-680 was also cloned into *BamHI* and *NotI* restriction sites of a modified pET-28b vector containing Smt3, a protein related to the small ubiquitin-like modifier (SUMO) family of proteins. This vector was a gift from the laboratory of Dr. Christopher Lima (Sloan-Kettering Institute for Cancer Research). Briefly, insert was cloned into the TOPO vector (Invitrogen) following PCR amplification with forward primer 5'-GAGAGCTCATG GCCGAGAACTTGCTG-3' and reverse primer 5'-GAAAGCTTTCAGTTACCCAG GATGCCTTGCTT-3'. The reading frame was verified by sequencing and plasmid DNA was isolated using GeneJet Plasmid Miniprep Kit (ThermoScientific). The vector and insert were digested with *BamHI* and *NotI*, and ligation followed as outlined above.

Plasmid DNA of CBP constructs was used to transform *E. coli* BL21 (DE3) cells (Novagen). Cells were grown overnight at 37°C in LB medium supplemented with 100 μg/mL kanamycin and 10 mL were used to inoculate each liter of Autoinduction Media containing the following: 1% tryptone, 0.5% yeast extract, 25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 2 mM $MgSO_4$, 54 mM glycerol, 2.8 mM glucose, and 5.6 mM α-lactose. Cells were grown at 18°C, protein expression was auto-induced, and growth continued overnight. Cells were harvested by centrifugation and washed with 50 mM sodium phosphate pH 8.0, 300 mM NaCl, 10 mM imidazole, and 10 mM β-mercaptoethanol. Following this wash step, cells were subsequently pelleted and stored at −80°C.

After thawing, cells were resuspended in lysis buffer: 50 mM sodium phosphate pH 8.0, 300 mM NaCl, 10 mM imidazole, and 10 mM β-mercaptoethanol, with one dissolved Complete EDTA-free protease inhibitor tablet (Roche), and lysed using microfluidization. Cellular debris was pelleted and removed by centrifugation on a Sorvall GSA rotor for 40 minutes at 12,000 rpm. The lysate was filtered through a 0.2 µm filter and loaded onto a HisTrap HP 5 mL column (GE Healthcare) pre-equilibrated with buffer containing 50 mM sodium phosphate pH 8.0, 300 mM NaCl, 10 mM imidazole, and 10 mM β-mercaptoethanol. Unbound proteins were washed for 5 column volumes with this buffer. Segmented imidazole gradient was used to elute protein with 30-100% of buffer containing 50 mM sodium phosphate pH 8.0, 300 mM NaCl, 250 mM imidazole, and 10 mM β-mercaptoethanol, to 30% over 20 column volumes, followed by holding step for 5 column volumes, with subsequent gradient to 100% over 20 column volumes in final step (Figure 64).

Fractions containing protein were pooled and dialyzed using 10 kD MWCO membrane (Pierce Snakeskin) against an excess of buffer containing 20 mM Tris-Cl pH 8.0, 300 mM NaCl, 5 mM β-mercaptoethanol, and Ulp1 protease was added to protein at 4°C with agitation overnight. The following morning, protein was dialyzed against an excess of buffer containing 20 mM Tris-Cl pH 8.0, 50 mM NaCl, and 5 mM β-mercaptoethanol. Subsequently, protein was loaded onto Source Q column (GE Healthcare) pre-equilibrated with filtered dialysis buffer to separate protein from tag. Elution buffer containing 20 mM Tris-Cl pH 8.0, 1 M NaCl, and 5 mM β-mercaptoethanol was used to elute uncleaved protein and bound tag from column. Unbound flow-through fractions containing tag-free protein were pooled and analyzed using SDS-PAGE analysis (Figure 65).

Protein fractions were analyzed by SDS-PAGE, and subsequently pooled and dialyzed against an excess of buffer A containing 100 mM MES pH 6.0, 50 mM NaCl,

0.05 mM $ZnCl_2$, and 10 mM β-mercaptoethanol for a total of 4 hours. The dialysis buffer was filtered through a 0.2 μm filter and used as the equilibration and binding buffer for the next step of purification. The protein was further purified using cation exchange chromatography over a Source S column (GE Healthcare). Unbound proteins were washed with 4 column volumes of buffer A with target protein eluted with a gradient to 100% of buffer B (100 mM MES pH 6.0, 1 M NaCl, 0.05 mM $ZnCl_2$, and 10 mM β-mercaptoethanol) over 20 column volumes (Figure 66). Fractions of protein containing ≥ 95% purity were collected and concentrated to 4-10 mg/mL, and stored at −80 °C.

7.2.2 Cloning and expression of CBP in insect cells

CBP (1-680) was cloned into the pBiEx1 vector using *NotI* and *SacI* as restriction sites using the same methods as outlined previously for cloning into pET vectors. Sf9 cells were grown to approximately 2.5×10^6 cells/mL in 2.5 L of serum-free Sf900IIISFM media (Invitrogen) without antibiotics. Plasmid DNA was added to serum-free media in a ratio of 1:2000 (420 μg DNA added to 84 mL of media) mixed by inversion, and filtered using 0.2 μm acetate syringe filters (Nalgene). A 250 μg/mL stock of cationic lipid, dimethyldioctadecylammonium bromide (DDAB), was prepared. DDAB was added to serum-free media in a 1:1 ratio (42 mL of each) and the liposome-DNA complexes were formed by the addition of DNA drop-wise to lipid, and the solution was mixed gently. Liposome-DNA complexes were incubated at room temperature for 20 minutes to allow binding of DNA to the liposome before adding to insect cells. The following day, 5% fetal bovine serum (FBS), and 1X stocks of penicillin and streptomycin were added to culture. Growth was continued at 27°C for sixty-five hours. The Sf9 cells were harvested by centrifugation at 1000 rpm for 15 minutes at 10°C and pellets stored at −80°C.

7.2.3 Purification from insect cells

Insect cell pellets were thawed in a beaker containing room temperature water and resuspended in buffer containing 50 mM sodium phosphate pH 8.0, 300 mM NaCl, 10 mM imidazole, 0.1% Triton X-100, 10 mM β-mercaptoethanol, with one dissolved Complete EDTA-free protease inhibitor tablet (Roche). Sonication was used to lyse cells at a 25% duty cycle for 1min with three second pulses and a one second pause in between pulses. Cellular debris was pelleted by centrifugation at 18,000 rpm at 4°C for 30 minutes using a Sorvall SS-34 rotor. The soluble cytoplasmic fraction was filtered with 0.2 μm SFCA-PF syringe filters (Corning) followed by another filtering step using 0.2 μm Stericup filters (Millipore). Protein lysate was incubated with 1 mL of a 50/50 slurry of Ni-NTA agarose beads (Qiagen) for 2 hours at 4°C with agitation. Beads were pelleted at low speed using centrifugation and washed with binding buffer (50 mM sodium phosphate pH 8.0, 300 mM NaCl, 10 mM imidazole, and 10 mM β-mercaptoethanol), for three 5 mL washes. Protein was eluted stepwise with 1 mL of buffer containing 50 mM sodium phosphate pH 8.0, 300 mM NaCl, and 10 mM β-mercaptoethanol with varying concentrations of imidazole (50 mM, 100 mM, 150 mM, or 250 mM) for a total of 12 washes. Fractions were analyzed by SDS-PAGE (Figure 67).

7.2.4 Screening of crystallization conditions

Recombinant protein was expressed from bacterial cells. CBP was screened for crystallization at concentrations between 4-10 mg/mL. Hanging drops consisting of 300 nL of protein were mixed with 300 nL of reservoir solution using the Mosquito (TTP Labtech). The following screens were used: Natrix Screen HT (Hampton Research), MPD (Qiagen), Crystal Screen I and II, Salt Rx HT, PEGRx HT (Hampton Research), and JSCG (Jena Bioscience). The total reservoir volume was 100 µL and trays were incubated at 20°C.

7.2.5 Electrophoretic Mobility Shift Assay

Electrophoretic mobility shift assay (EMSA) was performed with CREB, CBP, and CREM recombinant proteins and ^{32}P-labeled DNA fragments. Proteins were synthesized using the TNT coupled transcription and translation reticulocyte lysate system with T7 polymerase (Promega), according to the manufacturer's protocol. Reaction components (Table 7) were assembled and incubated at 30°C for 90 minutes. In order to assess CREB-CBP binding and CREM-CBP binding, a consensus cAMP response element (CRE) DNA binding site was generated using the following primers: sense 5'-CCTTGG CTGACGTCAGCCAAG-3', as well as antisense 5'-CTTGGCTGACGTCAGCCA AGG-3'. These oligonucleotides were γ^{32}P-end labeled with T4 polynucleotide kinase. For each EMSA, *in vitro* translated proteins were mixed with radiolabeled probe for 30 minutes at room temperature, along with deoxyinosine-deoxycytosine, salmon sperm, and binding buffer containing 50 mM potassium chloride, 20% glycerol, and 20 mm HEPES (pH 7.6-7.8). Binding was competed by addition of the unlabeled oligonucleotide in excess. Gel electrophoresis was used to separate each sample on a 5% nondenaturing acrylamide gel, followed by analysis using autoradiography.

Table 7. T_NT Reaction Components for EMSA

Reaction Components	**Amount Used**
T_NT Rabbit Reticulocyte Lysate	25 μL
T_NT Reaction Buffer	2 μL
T_NT RNA Polymerase	1 μL
Amino Acid Mixture, Minus Leucine, 1 mM	0.5 μL
Amino Acid Mixture, Minus Methionine, 1 mM	0.5 μL
RNasin Ribonculease Inhibitor (40 u/μL)	1 μL
DNA Templates (0.5 μg/μL)	2 μL
Transcend Biotin-Lysyl-tRNA	1 μL
Nuclease-free ddH_2O to final volume	50 μL

7.2.6 Limited proteolysis of CBP-CREM complex

Stocks of protein and protease were dissolved in the following buffer: Tris-Cl pH 8.0, 150 mM NaCl, and 10 mM BME. Reactions were prepared with CREM, CBP, CREM and CRE, CBP and CRE, or with a combination of all three to map binding interactions in the presence of trypsin. Equivalent amounts of CREM and CBP were used in each reaction (2.8 nanomoles), and trypsin was used at a final concentration of 1 μg/μL in a total reaction volume of 550 μL. $MgCl_2$ (2.5 mM) was added to all reactions containing equimolar amounts of the DNA duplex CRE. Following the addition of trypsin, reactions were allowed to proceed on ice, and samples were taken at 3 minutes, 60 minutes, and overnight. Following SDS-PAGE analysis (Figure 68) samples were transferred to PVDF, and subjected to N-terminal sequencing.

7.2.7 Gel Filtration of CBP and CREM

Gel filtration chromatography was used to characterize the binding interactions between CBP and CREM. CREM, CBP, and equimolar amounts of CBP and CREM were loaded onto a Superose 12 10/300 GL column (GE Healthcare) and eluted in one column volume with buffer containing 20 mM Tris-Cl pH 8.0, 150 mM NaCl, 10 mM β-mercaptoethanol (Figure 69). Fractions of CREM corresponding to the peak were concentrated to 10-12 mg/mL, and fractions of CBP were concentrated to 5-10 mg/mL for storage at −80°C.

Table 8. Expression systems utilized for production of CBP protein.

Construct	Expression System	Problem
pET42a-CBP302-680	T7	Precipitates easily
pET28a-CBP302-680	T7	Precipitates easily
pET42a-CBP1-680	T7	Precipitates easily
pET28a-CBP1-680	T7	Precipitates easily
pBiEx-3BS-CBP1-680	Insect	No expression
His6-Smt3-CBP1-680	T7	*No problem

*Construct was well-behaved.

7.2.7 Isothermal Calorimetry of CBP and CREM

The binding interactions of CBP and CREM were studied using isothermal titration calorimetry (ITC) using a VP-ITC instrument (Microcal Inc.). Both proteins were dialyzed overnight against 1 L of buffer containing: 25 mM HEPES pH 7.5, 300 mM NaCl, 5 mM $MgCl_2$, and 1 mM TCEP. Subsequently, dialysis buffer was filtered using a 0.2 μm filter and used in the binding experiments. All solutions were degassed immediately before use. CBP was diluted to 75 μM, and CREM was diluted to 350 μM with dialysis buffer. In each case, these calculations were based on the molecular weight of the monomer following assessment of protein concentration. In the sample cell, 1.4 mL of CBP was titrated with twenty-four 10 μL injections of CREM, after an initial 2 μL injection. CREM was added at 3.5 minute intervals with stirring at 300 rpm. Experiments were performed at 27°C, using a reference power of 5 μCal/second. The MicroCal ORIGIN software package 5.0 was used to analyze data. The enthalpy of binding was determined based on the heat released.

7.3 Results and Discussion

In these studies we attempted to use a construct of CBP encompassing multiple domains, and a construct of full-length CREB in order to identify local protein interactions as well as determine the long-range effects of protein-protein interactions between these two proteins. Binary complexes of CREB:CBP or ternary complexes of CREB:CRE:CBP would reveal the necessary binding interactions that must occur for complex assembly and for the induction of CREB transactivation. Structural studies would have allowed us to dissect the role of each partner in the assembly of the complex and assist us in determining which regions interact to mediate CREB-dependent gene transcription.

Despite the use of several different constructs and expression systems (Table 8), X-ray quality crystals were not obtained. The soluble monomeric tag Smt3/SUMO has been shown to increase protein solubility [242], and we observed an increase in overall yield with this construct in comparison to all other CBP constructs. However no crystallization screens were performed with this SUMO-tagged construct.

In addition, we employed several different methods to assess the binding interactions of CBP and CREM. EMSA, limited proteolytic cleavage assay, gel filtration, and ITC were all utilized. For the experiments with EMSA, there were no gel shifts indicating a lack of interaction between CBP and CREM. The results from the proteolytic cleavage assay followed by N-terminal sequencing, and from the gel filtration experiments were promising. These results indicated that CBP delays the proteolytic cleavage of CREM. In its absence CREM was more accessible, and the N-terminal sequencing results demonstrated that C-terminal arginine residues present in CREM were available for digestion by trypsin. CREM contains a kinase inducible domain and it seemed to bind CBP through this region. This interaction had a protective effect against digestion. Additionally, in the gel filtration studies, there was a shift in the peak when CREM and CBP were combined in equimolar amounts in comparison to the peak observed for either protein alone.

By employing ITC, we attempted to map the binding site of CBP and CREM, as well as confirm the binding interactions we observed during gel filtration and proteolytic cleavage experiments with trypsin. We had hoped to gain information about the binding affinity and to detect the presence of multiple or different binding sites between CBP and CREM. However, ITC experiments only indicated heats of dilution rather than the enthalpy of binding (data not shown).

Additionally we used the XtalPred web server to predict the crystallizability of CBP [229]. A prediction was made based on a comparison of different protein features with those present in a target database [230]. CBP belongs to the class with the highest degree of difficulty and the results are summarized (Figure 70). It contains a long disordered region as well as other features that make crystallization difficult.

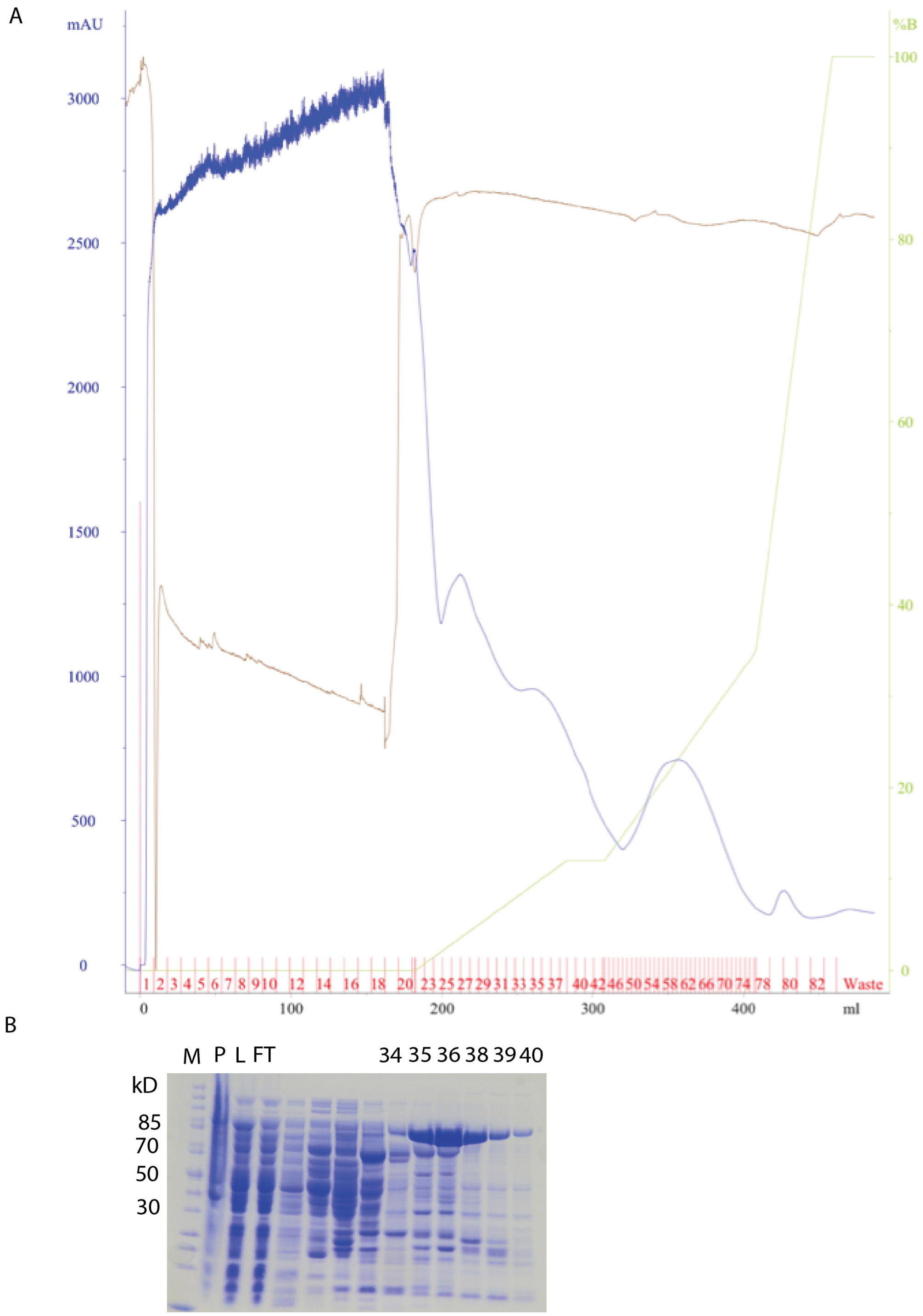


Figure 64. Ni^{2+}-affinity chromatography of CBP.
(A) Chromatogram of CBP purification. The blue line denotes absorbance at 280 nm in milliabsorbance units on the left and the green line represents gradient of buffer used for elution with concentration listed on right axis. Pink line denotes sample injection. (B) SDS-PAGE gel of purification stained with Coomassie blue stain. M refers to molecular weight ladder in kilodaltons, P refers to insoluble pellet, L to sample loaded onto column, FT to unbound protein. Fractions pooled for dialysis are shown.

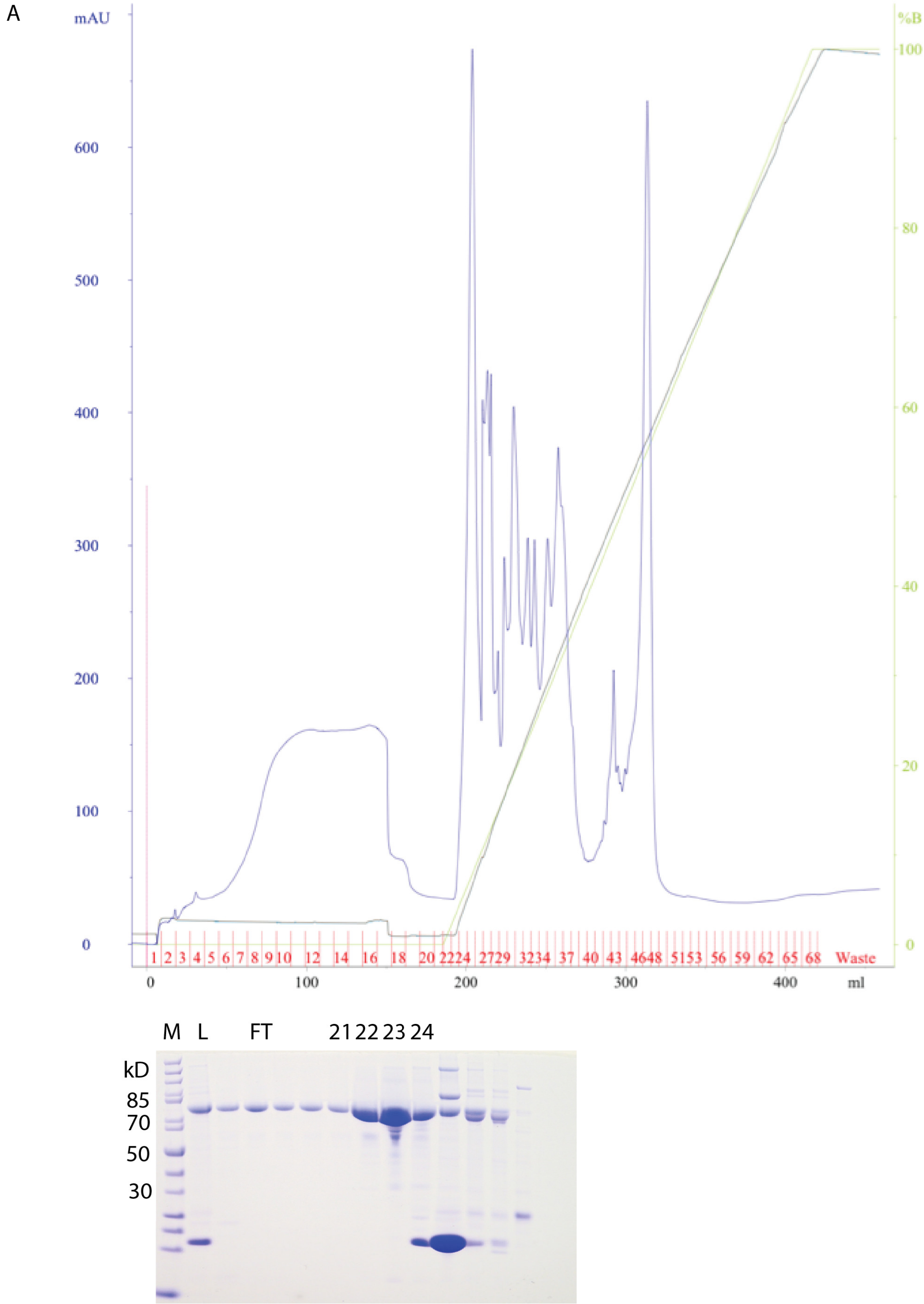


Figure 65. Anion exchange chromatography of CBP after Ulp1 cleavage.
(A) Chromatogram of second step of CBP purification. The blue line denotes absorbance at 280 nm in milliabsorbance units on the left and the green line represents gradient of buffer used for elution with concentration listed on right axis. Pink line denotes sample injection. (B) SDS-PAGE gel of purification stained with Coomassie blue stain. M refers to molecular weight ladder in kilodaltons and L to sample loaded onto column, FT refers to unbound/tag-free protein. Fractions 5-23 were pooled and dialyzed for next step.

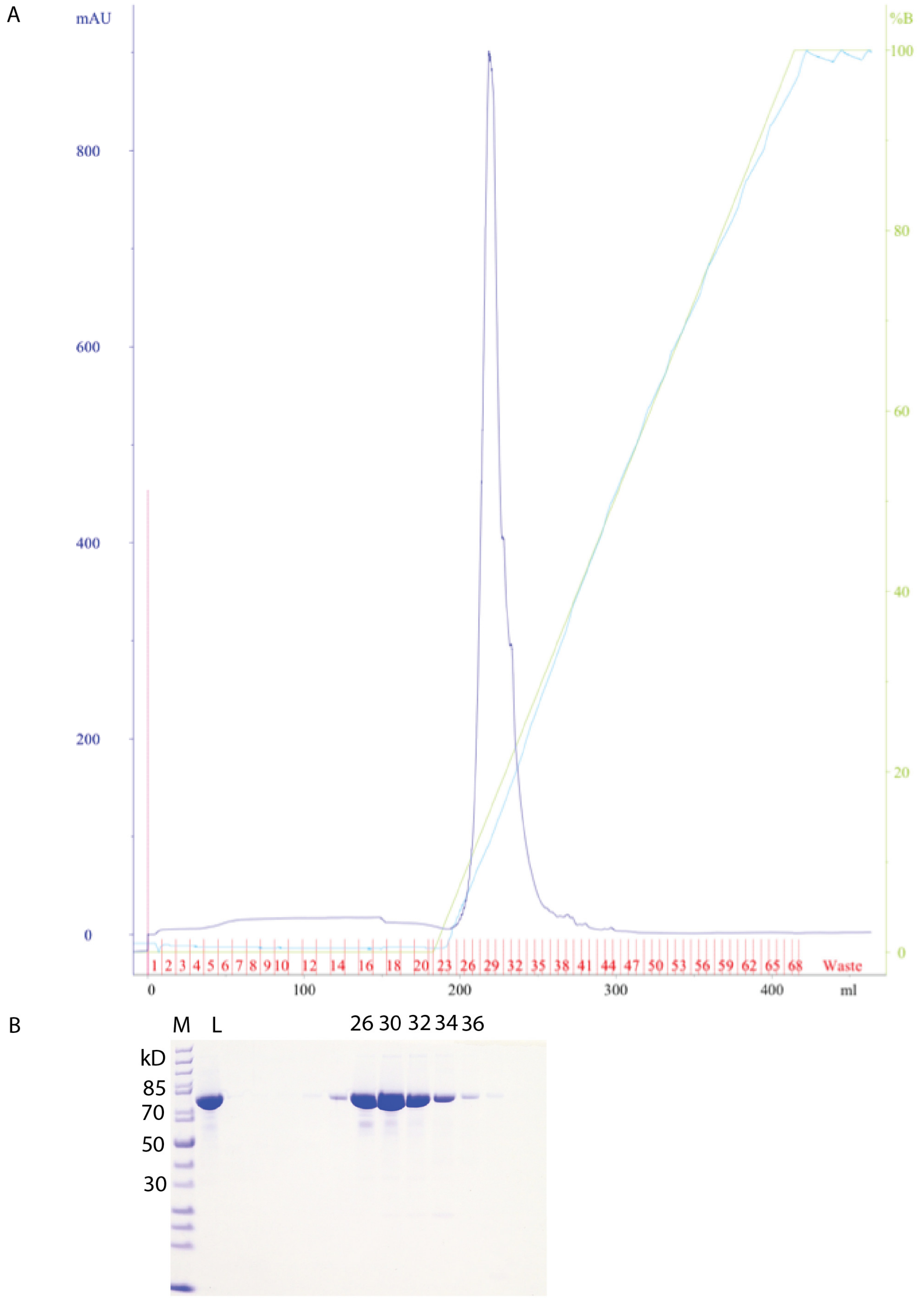


Figure 66. Cation exchange chromatography of CBP.
(A) Chromatogram of last step of CBP purification. The blue line denotes absorbance at 280 nm in milliabsorbance units and the green line represents gradient of buffer used for elution. Pink line denotes sample injection. (B) SDS-PAGE gel of purification stained with Coomassie blue stain. M refers to molecular weight ladder in kilodaltons and L to sample loaded onto column. Fractions corresponding to peak are shown.

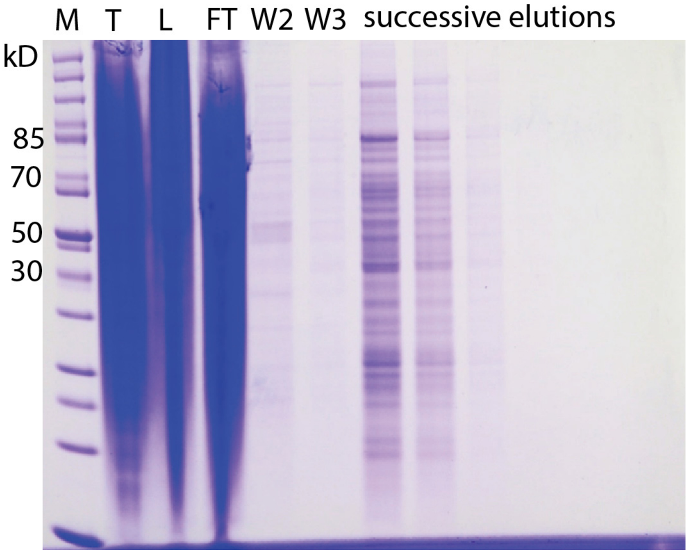


Figure 67. Ni^{2+}-affinity chromatograpy of CBP using Ni-NTA agarose beads.
SDS-PAGE gel displaying purification of CBP from insect cell lysate. The gel was stained with Coomassie blue stain. M refers to molecular weight ladder in kilodaltons, T refers to total lysate, L refers to sample loaded onto column, FT to unbound protein, W refers to wash steps, and successive elution steps are also displayed.

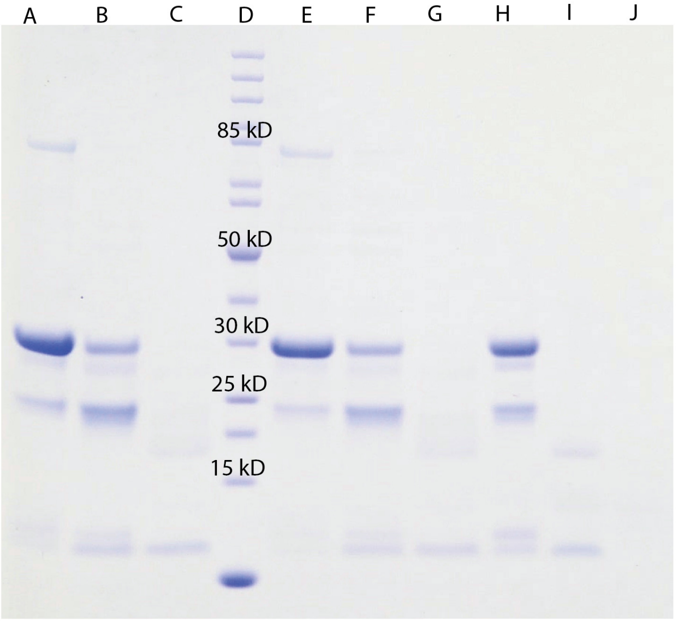


Figure 68. CBP protects CREM from proteolytic cleavage.
SDS PAGE analysis after incubation with trypsin. (A) CBP and CREM after 3 minutes; (B) CBP and CREM after 60 minutes; (C) CBP and CREM overnight; (D) molecular weight, (E) CBP, CREM, CRE after 3 minutes; (F) CBP, CREM, CRE after 60 minutes; (G) CBP, CREM, CRE overnight; (H) CREM and CRE after 3 minutes; (I) CREM and CRE after 60 minutes; (J) CREM and CRE overnight.

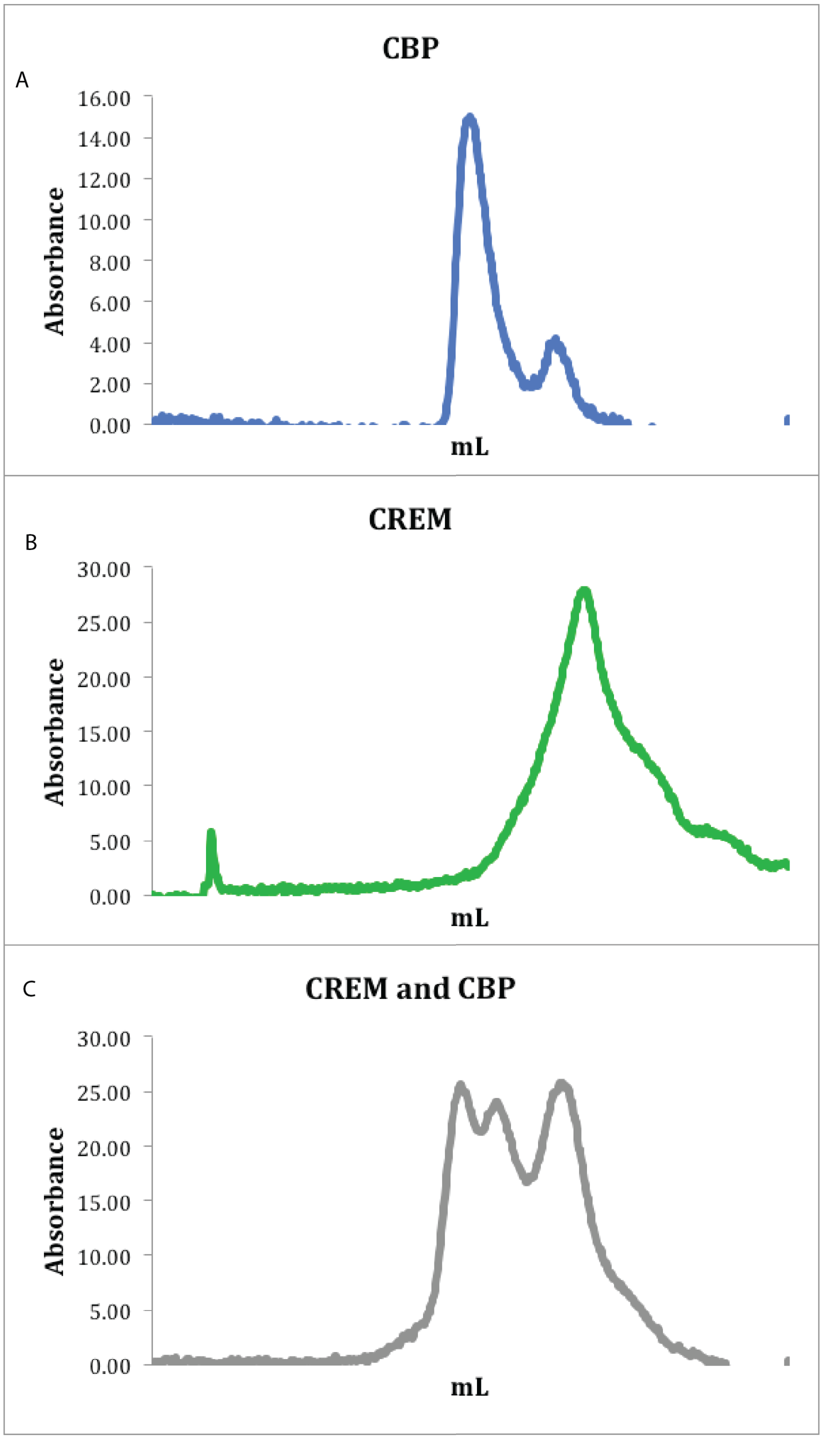
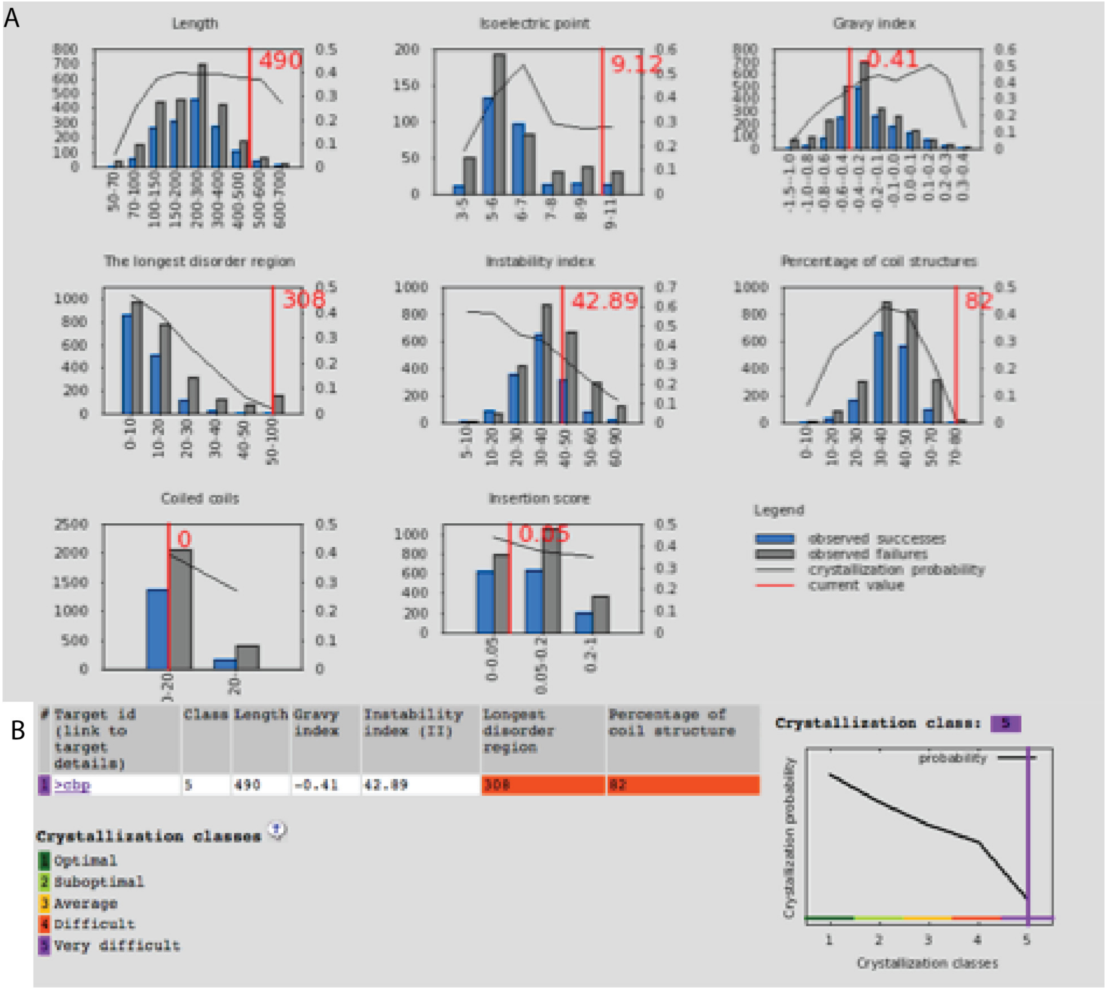


Figure 69. CBP and CREM interact.
Peaks shift upon interaction. (A) Gel filtration chromatography of CBP. (B) Gel filtration chromatography of CREM. (C) Gel filtration chromatography of CREM and CBP.



#	Target id (link to target details)	Class	Length	Gravy index	Instability index (II)	Longest disorder region	Percentage of coil structure
1	>cbp	5	490	-0.41	42.89	308	82

Crystallization classes

- 1 Optimal
- 2 Suboptimal
- 3 Average
- 4 Difficult
- 5 Very difficult

Figure 70. Crystallization probability analysis of CBP.
(A) Histograms representing comparison of target features with distribution probabilities obtained for: length, isoelectric point, gravity index, length of disordered regions, instability, percentage of coil structures, and number of coiled coils displayed for CBP. (B) Crystallization class with length, gravity index, instability index, and longest disordered region listed for CBP. Legend denotes degree of difficulty by color.

List of References

1 Chen, R. and Jeong, S.-S. (2000) Functional prediction: Identification of protein orthologs and paralogs. *Protein Science* 9, 2344-2353.

2 Tipton, P.A. and Beecher, B.S. (1994) Tartrate Dehydrogenase, a New Member of the Family of Metal-Dependent Decarboxylating R-Hydroxyacid Dehydrogenases. *Archives of Biochemistry and Biophysics* 313, 15-21.

3 Hurley, J.H.*, et al.* (1991) Catalytic mechanism of NADP(+)-dependent isocitrate dehydrogenase: implications from the structures of magnesium-isocitrate and NADP+ complexes. *Biochemistry* 30, 8671-8678.

4 Imada, K.*, et al.* (1991) Three-dimensional structure of a highly thermostable enzyme, 3-isopropylmalate dehydrogenase of Thermus thermophilus at 2.2 Å resolution. *Journal of Molecular Biology* 222, 725-738.

5 Kohn, L.D.*, et al.* (1968) Tartaric Acid Metabolism: V. CRYSTALLINE TARTRATE DEHYDROGENASE. *Journal of Biological Chemistry* 243, 2479-2485.

6 Strassman, M. and Ceci, L.N. (1965) Enzymatic Formation of α-Ketoadipic Acid from Homoisocitric Acid. *Journal of Biological Chemistry* 240, 4357-4361.

7 Chen, R.D. and Gadal, P. (1990) DO THE MITOCHONDRIA PROVIDE THE 2-OXOGLUTARATE NEEDED FOR GLUTAMATE SYNTHESIS IN HIGHER-PLANT CHLOROPLASTS. *Plant Physiology and Biochemistry* 28, 141-145

8 Nekrutenko, A.*, et al.* (1998) Cytosolic isocitrate dehydrogenase in humans, mice, and voles and phylogenetic analysis of the enzyme family. *Molecular Biology and Evolution* 15, 1674-1684.

9 Henke, B.*, et al.* (1998) IDP3 Encodes a Peroxisomal NADP-dependent Isocitrate Dehydrogenase Required for the β-Oxidation of Unsaturated Fatty Acids. *Journal of Biological Chemistry* 273, 3702-3711.

10 Frederiks, W.M.*, et al.* (2007) NADPH Production by the Pentose Phosphate Pathway in the Zona Fasciculata of Rat Adrenal Gland. *Journal of Histochemistry & Cytochemistry* 55, 975-980.

11 Marino, D.*, et al.* (2007) NADPH recycling systems in oxidative stressed pea nodules: a key role for the NADP+-dependent isocitrate dehydrogenase. *Planta* 225, 413-421.

12 LaPorte, D.C. and Koshland, D.E., Jr. (1982) A protein with kinase and phosphatase activities involved in regulation of tricarboxylic acid cycle. *Nature* 300, 458-460.

13 Garnak, M. and Reeves, H.C. (1979) Phosphorylation of Isocitrate dehydrogenase of Escherichia coli. *Science* 203, 1111-1112.

14 LaPorte, D.C. and Koshland, D.E. (1983) Phosphorylation of isocitrate dehydrogenase as a demonstration of enhanced sensitivity in covalent regulation. *Nature* 305, 286-290.

15 Laporte, D.C. (1993) The isocitrate dehydrogenase phosphorylation cycle: Regulation and enzymology. *Journal of Cellular Biochemistry* 51, 14-18.

16 Chen, R.F. and Plaut, G.W.E. (1963) Activation and Inhibition of DPN-linked Isocitrate Dehydrogenase of Heart by Certain Nucleotides*. *Biochemistry* 2, 1023-1032.

17 Hathaway, J.A. and Atkinson, D.E. (1963) The Effect of Adenylic Acid on Yeast Nicotinamide Adenine Dinucleotide Isocitrate Dehydrogenase, a Possible Metabolic Control Mechanism. *Journal of Biological Chemistry* 238, 2875-2881.

18 Xu, X.*, et al.* (2004) Structures of human cytosolic NADP-dependent isocitrate dehydrogenase reveal a novel self-regulatory mechanism of activity. *The Journal of biological chemistry* 279, 33946-33957.

19 Grzeschik, K.H. (1976) Assignment of a gene for human mitochondrial isocitrate dehydrogenase (ICD-M, EC 1.1.1.41) to chromosome 15. *Human genetics* 34, 23-28.

20 Huh, T.L.*, et al.* (1996) Assignment of the human mitochondrial NAD+ -specific isocitrate dehydrogenase alpha subunit (IDH3A) gene to 15q25.1-->q25.2by in situ hybridization. *Genomics* 32, 295-296.

21 Kim, Y.O.*, et al.* (1999) Assignment of mitochondrial NAD(+)-specific isocitrate dehydrogenase beta subunit gene (IDH3B) to human chromosome band 20p13 by in situ hybridization and radiation hybrid mapping. *Cytogenetics and cell genetics* 86, 240-241.

22 Narahara, K.*, et al.* (1985) Probable assignment of soluble isocitrate dehydrogenase (IDH1) to 2q33.3. *Human genetics* 71, 37-40.

23 Ramachandran, N. and Colman, R.F. (1980) Chemical characterization of distinct subunits of pig heart DPN-specific isocitrate dehydrogenase. *The Journal of biological chemistry* 255, 8859-8864.

24 Soundar, S.*, et al.* (2006) Identification of Mn2+-binding Aspartates from α, β, and γ Subunits of Human NAD-dependent Isocitrate Dehydrogenase. *Journal of Biological Chemistry* 281, 21073-21081.

25 Geisbrecht, B.V. and Gould, S.J. (1999) The human PICD gene encodes a cytoplasmic and peroxisomal NADP(+)-dependent isocitrate dehydrogenase. *The Journal of biological chemistry* 274, 30527-30533.

26 Koh, H.-J.*, et al.* (2004) Cytosolic NADP+-dependent Isocitrate Dehydrogenase Plays a Key Role in Lipid Metabolism. *Journal of Biological Chemistry* 279, 39968-39974.

27 Hausinger, R.P. (2004) Fe(II)/α -Ketoglutarate-Dependent Hydroxylases and Related Enzymes. *Critical Reviews in Biochemistry & Molecular Biology* 39, 21-68.

28 Kovaleva, E.G. and Lipscomb, J.D. (2008) Versatility of biological non-heme Fe(II) centers in oxygen activation reactions. *Nat Chem Biol* 4, 186-193.

29 Loenarz, C. and Schofield, C.J. (2008) Expanding chemical biology of 2-oxoglutarate oxygenases. *Nat Chem Biol* 4, 152-156.

30 Schofield, C.J. and Zhang, Z. (1999) Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes. *Current Opinion in Structural Biology* 9, 722-731.

31 Jo, S.-H.*, et al.* (2001) Control of Mitochondrial Redox Balance and Cellular Defense against Oxidative Damage by Mitochondrial NADP+-dependent Isocitrate Dehydrogenase. *Journal of Biological Chemistry* 276, 16168-16176.

32 Kim, S.Y. and Park, J.-W. (2003) Cellular Defense against Singlet Oxygen-induced Oxidative Damage by Cytosolic NADP + -dependent Isocitrate Dehydrogenase. *Free Radical Research* 37, 309-316.
33 Lee, S.M.*, et al.* (2002) Cytosolic NADP+-dependent isocitrate dehydrogenase status modulates oxidative damage to cells. *Free Radical Biology and Medicine* 32, 1185-1196.
34 Kaimul, A.M.*, et al.* (2007) Thioredoxin and thioredoxin-binding protein-2 in cancer and metabolic syndrome. *Free Radical Biology and Medicine* 43, 861-868.
35 Arnér, E.S.J. and Holmgren, A. (2000) Physiological functions of thioredoxin and thioredoxin reductase. *European Journal of Biochemistry* 267, 6102-6109.
36 Alp, P.R.*, et al.* (1976) Activities of citrate synthase and NAD+-linked and NADP+-linked isocitrate dehydrogenase in muscle from vertebrates and invertebrates. *The Biochemical journal* 154, 689-700.
37 Hoek, J.B. and Rydstrom, J. (1988) Physiological roles of nicotinamide nucleotide transhydrogenase. *The Biochemical journal* 254, 1-10.
38 Hartong, D.T.*, et al.* (2008) Insights from retinitis pigmentosa into the roles of isocitrate dehydrogenases in the Krebs cycle. *Nature genetics* 40, 1230-1234.
39 Rossmann, M.G.*, et al.* (1974) Chemical and biological evolution of nucleotide-binding protein. *Nature* 250, 194-199.
40 Rao, S.T. and Rossmann, M.G. (1973) Comparison of super-secondary structures in proteins. *Journal of Molecular Biology* 76, 241-256.
41 Wierenga, R.K.*, et al.* (1986) Prediction of the occurrence of the ADP-binding βαβ-fold in proteins, using an amino acid sequence fingerprint. *Journal of Molecular Biology* 187, 101-107.
42 Kleiger, G. and Eisenberg, D. (2002) GXXXG and GXXXA Motifs Stabilize FAD and NAD(P)-binding Rossmann Folds Through Cα–H···O Hydrogen Bonds and van der Waals Interactions. *Journal of Molecular Biology* 323, 69-76.
43 Dang, L.*, et al.* (2010) Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* 465, 966.
44 Ceccarelli, C.*, et al.* (2002) Crystal Structure of Porcine Mitochondrial NADP+-dependent Isocitrate Dehydrogenase Complexed with Mn2+ and Isocitrate: INSIGHTS INTO THE ENZYME MECHANISM. *Journal of Biological Chemistry* 277, 43454-43462.
45 Siebert, G.*, et al.* (1957) THE ENZYMATIC PROPERTIES OF ISOCITRIC DEHYDROGENASE. *Journal of Biological Chemistry* 226, 977-991.
46 Dalziel, K. and Londesborough, J.C. (1968) The mechanisms of reductive carboxylation reactions. Carbon dioxide or bicarbonate as substrate of nicotinamide-adenine dinucleotide phosphate-linked isocitrate dehydrogenase and malic enzyme. *The Biochemical journal* 110, 223-230.
47 Grodsky, N.B.*, et al.* (2000) Evaluation by site-directed mutagenesis of aspartic acid residues in the metal site of pig heart NADP-dependent isocitrate dehydrogenase. *Biochemistry* 39, 2193-2200.
48 Kim, T.-K.*, et al.* (2003) Critical Role of Lys212 and Tyr140 in Porcine NADP-dependent Isocitrate Dehydrogenase. *Journal of Biological Chemistry* 278, 49323-49331.

49 Bolduc, J.M., *et al.* (1995) Mutagenesis and Laue structures of enzyme intermediates: isocitrate dehydrogenase. *Science* 268, 1312-1318.

50 Kim, T.-K. and Colman, R.F. (2005) Ser95, Asn97, and Thr78 are important for the catalytic function of porcine NADP-dependent isocitrate dehydrogenase. *Protein Science* 14, 140-147.

51 Lee, M.E., *et al.* (1995) Mutational analysis of the catalytic residues lysine 230 and tyrosine 160 in the NADP(+)-dependent isocitrate dehydrogenase from Escherichia coli. *Biochemistry* 34, 378-384.

52 Warburg, O. (1956) On respiratory impairment in cancer cells. *Science* 124, 269-270.

53 Warburg, O. (1956) On the origin of cancer cells. *Science* 123, 309-314.

54 Warburg, O., *et al.* (1927) The Metabolism of Tumors in the Body. *The Journal of general physiology* 8, 519-530.

55 Vander Heiden, M.G., *et al.* (2009) Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* 324, 1029-1033.

56 Patronas, N.J., *et al.* (1985) Prediction of survival in glioma patients by means of positron emission tomography. *Journal of neurosurgery* 62, 816-822.

57 Tralins, K.S., *et al.* (2002) Volumetric analysis of 18F-FDG PET in glioblastoma multiforme: prognostic information and possible role in definition of target volumes in radiation dose escalation. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine* 43, 1667-1673.

58 Levine, A.J. and Puzio-Kuter, A.M. (2010) The control of the metabolic switch in cancers by oncogenes and tumor suppressor genes. *Science* 330, 1340-1344.

59 Kroemer, G. and Pouyssegur, J. (2008) Tumor Cell Metabolism: Cancer's Achilles' Heel. *Cancer cell* 13, 472-482.

60 Vogelstein, B., *et al.* (2013) Cancer Genome Landscapes. *Science* 339, 1546-1558.

61 Knudson, A.G., Jr. (1971) Mutation and cancer: statistical study of retinoblastoma. *Proceedings of the National Academy of Sciences of the United States of America* 68, 820-823.

62 Vogelstein, B. and Kinzler, K.W. (1993) The multistep nature of cancer. *Trends in Genetics* 9, 138-141.

63 Dang, C.V. and Semenza, G.L. (1999) Oncogenic alterations of metabolism. *Trends in biochemical sciences* 24, 68-72.

64 Astuti, D., *et al.* (2001) Germline SDHD mutation in familial phaeochromocytoma. *Lancet* 357, 1181-1182.

65 Astuti, D., *et al.* (2001) Gene mutations in the succinate dehydrogenase subunit SDHB cause susceptibility to familial pheochromocytoma and to familial paraganglioma. *American journal of human genetics* 69, 49-54.

66 Niemann, S. and Muller, U. (2000) Mutations in SDHC cause autosomal dominant paraganglioma, type 3. *Nature genetics* 26, 268-270.

67 Korpershoek, E., *et al.* (2011) SDHA immunohistochemistry detects germline SDHA gene mutations in apparently sporadic paragangliomas and pheochromocytomas. *The Journal of clinical endocrinology and metabolism* 96, E1472-1476.

68 Pagnamenta, A.T., *et al.* (2006) Phenotypic variability of mitochondrial disease caused by a nuclear mutation in complex II. *Molecular Genetics and Metabolism* 89, 214-221.

69 Pollard, P.J., *et al.* (2005) Accumulation of Krebs cycle intermediates and over-expression of HIF1alpha in tumours which result from germline FH and SDH mutations. *Human molecular genetics* 14, 2231-2239.

70 Tomlinson, I.P., *et al.* (2002) Germline mutations in FH predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer. *Nature genetics* 30, 406-410.

71 Vanharanta, S., *et al.* (2006) Distinct expression profile in fumarate-hydratase-deficient uterine fibroids. *Human molecular genetics* 15, 97-103.

72 Isaacs, J.S., *et al.* (2005) HIF overexpression correlates with biallelic loss of fumarate hydratase in renal cancer: novel role of fumarate in regulation of HIF stability. *Cancer cell* 8, 143-153.

73 Wang, G.L., *et al.* (1995) Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension. *Proceedings of the National Academy of Sciences of the United States of America* 92, 5510-5514.

74 Kaelin, W.G., Jr. and Ratcliffe, P.J. (2008) Oxygen sensing by metazoans: the central role of the HIF hydroxylase pathway. *Molecular cell* 30, 393-402.

75 Maxwell, P.H., *et al.* (2001) Activation of the HIF pathway in cancer. *Current opinion in genetics & development* 11, 293-299.

76 Tennant, D.A. and Gottlieb, E. (2010) HIF prolyl hydroxylase-3 mediates alpha-ketoglutarate-induced apoptosis and tumor suppression. *J Mol Med (Berl)* 88, 839-849.

77 Ke, Q. and Costa, M. (2006) Hypoxia-inducible factor-1 (HIF-1). *Molecular pharmacology* 70, 1469-1480.

78 Ostrom, Q.T., *et al.* (2013) CBTRUS Statistical Report: Primary Brain and Central Nervous System Tumors Diagnosed in the United States in 2006-2010. *Neuro-oncology* 15, ii1-ii56.

79 Furnari, F.B., *et al.* (2007) Malignant astrocytic glioma: genetics, biology, and paths to treatment. *Genes & development* 21, 2683-2710.

80 Louis, D.N., *et al.* (2007) The 2007 WHO classification of tumours of the central nervous system. *Acta neuropathologica* 114, 97-109.

81 Brem, S.S., *et al.* (1974) Angiogenesis in brain tumors: a quantitative histologic study. *Surgical forum* 25, 462-464.

82 Folkman, J. (2007) Angiogenesis: an organizing principle for drug discovery? *Nature reviews. Drug discovery* 6, 273-286.

83 Leon, S.P., *et al.* (1996) Microvessel density is a prognostic indicator for patients with astroglial brain tumors. *Cancer* 77, 362-372.

84 Birlik, B., *et al.* (2006) Tumour vascularity is of prognostic significance in adult, but not paediatric astrocytomas. *Neuropathology and applied neurobiology* 32, 532-538.

85 Takano, S., *et al.* (1996) Concentration of vascular endothelial growth factor in the serum and tumor tissue of brain tumor patients. *Cancer research* 56, 2185-2190.

86 Ferrara, N. (2002) Role of vascular endothelial growth factor in physiologic and pathologic angiogenesis: therapeutic implications. *Semin Oncol* 29, 10-14.

87 Goussia, A.C.*, et al.* (2000) Cytogenetic and molecular genetic abnormalities in primitive neuroectodermal tumors of the central nervous system. *Anticancer research* 20, 65-73.

88 Gerstner, E.R.*, et al.* (2009) VEGF inhibitors in the treatment of cerebral edema in patients with brain cancer. *Nature reviews. Clinical oncology* 6, 229-236.

89 Stevens, G.H.J. (2010) Brain Tumors:Meningiomas and Gliomas. http://www.clevelandclinicmeded.com/medicalpubs/diseasemanagement/hematology-oncology/brain-tumors/.

90 Kleihues, P. and Ohgaki, H. (1999) Primary and secondary glioblastomas: from concept to clinical diagnosis. *Neuro-oncology* 1, 44-51.

91 Maher, E.A.*, et al.* (2006) Marked genomic differences characterize primary and secondary glioblastoma subtypes and identify two distinct molecular and clinical secondary glioblastoma entities. *Cancer research* 66, 11502-11513.

92 Nakamura, M.*, et al.* (2005) Frequent LOH on 22q12.3 and TIMP-3 inactivation occur in the progression to secondary glioblastomas. *Laboratory investigation; a journal of technical methods and pathology* 85, 165-175.

93 Ohgaki, H.*, et al.* (2004) Genetic pathways to glioblastoma: a population-based study. *Cancer research* 64, 6892-6899.

94 Noushmehr, H.*, et al.* (2010) Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer cell* 17, 510-522.

95 Ohgaki, H. and Kleihues, P. (2007) Genetic pathways to primary and secondary glioblastoma. *The American journal of pathology* 170, 1445-1453.

96 Uhlmann, K.*, et al.* (2003) Distinct methylation profiles of glioma subtypes. *International Journal of Cancer* 106, 52-59.

97 Dong, S.M.*, et al.* (2001) Concurrent hypermethylation of multiple genes is associated with grade of oligodendroglial tumors. *Journal of neuropathology and experimental neurology* 60, 808-816.

98 Stupp, R.*, et al.* (2006) Changing Paradigms—An Update on the Multidisciplinary Management of Malignant Glioma. *The Oncologist* 11, 165-180.

99 Ino, Y.*, et al.* (2001) Molecular subtypes of anaplastic oligodendroglioma: implications for patient management at diagnosis. *Clinical cancer research : an official journal of the American Association for Cancer Research* 7, 839-845.

100 Stupp, R.*, et al.* (2005) Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *The New England journal of medicine* 352, 987-996.

101 Grossman, S.A. and Batara, J.F. (2004) Current management of glioblastoma multiforme. *Seminars in Oncology* 31, 635-644.

102 Friedman, H.S.*, et al.* (1998) DNA mismatch repair and O6-alkylguanine-DNA alkyltransferase analysis and response to Temodal in newly diagnosed malignant glioma. *Journal of Clinical Oncology* 16, 3851-3857.

103 Hotta, T.*, et al.* (1994) O6-alkylguanine-DNA alkyltransferase activity of human malignant glioma and its clinical implications. *Journal of neuro-oncology* 21, 135-140.

104 Sjoblom, T., *et al.* (2006) The consensus coding sequences of human breast and colorectal cancers. *Science* 314, 268-274.

105 Parsons, D.W., *et al.* (2008) An integrated genomic analysis of human glioblastoma multiforme. *Science* 321, 1807-1812.

106 Amary, M.F., *et al.* (2011) IDH1 and IDH2 mutations are frequent events in central chondrosarcoma and central and periosteal chondromas but not in other mesenchymal tumours. *The Journal of pathology* 224, 334-343.

107 Kang, M.R., *et al.* (2009) Mutational analysis of IDH1 codon 132 in glioblastomas and other common cancers. *International journal of cancer. Journal international du cancer* 125, 353-355.

108 Kipp, B.R., *et al.* (2012) Isocitrate dehydrogenase 1 and 2 mutations in cholangiocarcinoma. *Human pathology* 43, 1552-1558.

109 Pansuriya, T.C., *et al.* (2011) Somatic mosaic IDH1 and IDH2 mutations are associated with enchondroma and spindle cell hemangioma in Ollier disease and Maffucci syndrome. *Nature genetics* 43, 1256-1261.

110 Kosmider, O., *et al.* (2010) Mutations of IDH1 and IDH2 genes in early and accelerated phases of myelodysplastic syndromes and MDS/myeloproliferative neoplasms. *Leukemia* 24, 1094-1096.

111 Murugan, A.K., *et al.* (2010) Identification and functional characterization of isocitrate dehydrogenase 1 (IDH1) mutations in thyroid cancer. *Biochemical and Biophysical Research Communications* 393, 555-559.

112 Gaal, J., *et al.* (2010) Isocitrate Dehydrogenase Mutations Are Rare in Pheochromocytomas and Paragangliomas. *Journal of Clinical Endocrinology & Metabolism* 95, 1274-1278.

113 Balss, J., *et al.* (2008) Analysis of the IDH1 codon 132 mutation in brain tumors. *Acta neuropathologica* 116, 597-602.

114 Bleeker, F.E., *et al.* (2009) IDH1 mutations at residue p.R132 (IDH1R132) occur frequently in high-grade gliomas but not in other solid tumors. *Human Mutation* 30, 7-11.

115 Green, A. and Beer, P. (2010) Somatic mutations of IDH1 and IDH2 in the leukemic transformation of myeloproliferative neoplasms. *The New England journal of medicine* 362, 369-370.

116 Hartmann, C., *et al.* (2009) Type and frequency of IDH1 and IDH2 mutations are related to astrocytic and oligodendroglial differentiation and age: a study of 1,010 diffuse gliomas. *Acta neuropathologica* 118, 469-474.

117 Ichimura, K., *et al.* (2009) IDH1 mutations are present in the majority of common adult gliomas but rare in primary glioblastomas. *Neuro-oncology* 11, 341-347.

118 Mardis, E.R., *et al.* (2009) Recurring mutations found by sequencing an acute myeloid leukemia genome. *The New England journal of medicine* 361, 1058-1066.

119 Yan, H., *et al.* (2009) IDH1 and IDH2 mutations in gliomas. *The New England journal of medicine* 360, 765-773.

120 Rakheja, D., *et al.* (2012) IDH mutations in acute myeloid leukemia. *Human pathology* 43, 1541-1551.

121 Dang, L., *et al.* (2010) IDH mutations in glioma and acute myeloid leukemia. *Trends in molecular medicine* 16, 387-397.

122 Gross, S., *et al.* (2010) Cancer-associated metabolite 2-hydroxyglutarate accumulates in acute myelogenous leukemia with isocitrate dehydrogenase 1 and 2 mutations. *The Journal of experimental medicine* 207, 339-344.
123 Marcucci, G., *et al.* (2010) IDH1 and IDH2 Gene Mutations Identify Novel Molecular Subsets Within De Novo Cytogenetically Normal Acute Myeloid Leukemia: A Cancer and Leukemia Group B Study. *Journal of Clinical Oncology* 28, 2348-2355.
124 Ward, P.S., *et al.* (2010) The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting alpha-ketoglutarate to 2-hydroxyglutarate. *Cancer cell* 17, 225-234.
125 Schaap, F.G., *et al.* (2013) Mutations in the isocitrate dehydrogenase genes IDH1 and IDH2 in tumors. *Advances in anatomic pathology* 20, 32-38.
126 Watanabe, T., *et al.* (2009) IDH1 mutations are early events in the development of astrocytomas and oligodendrogliomas. *The American journal of pathology* 174, 1149-1153.
127 Labussière, M., *et al.* (2010) All the 1p19q codeleted gliomas are mutated on IDH1 or IDH2. *Neurology* 74, 1886-1890.
128 Bujko, M., *et al.* (2010) Prognostic value of IDH1 mutations identified with PCR-RFLP assay in glioblastoma patients. *Molecular diagnosis & therapy* 14, 163-169.
129 Metellus, P., *et al.* (2010) Absence of IDH mutation identifies a novel radiologic and molecular subtype of WHO grade II gliomas with dismal prognosis. *Acta neuropathologica* 120, 719-729.
130 Nobusawa, S., *et al.* (2009) IDH1 Mutations as Molecular Signature and Predictive Factor of Secondary Glioblastomas. *Clinical Cancer Research* 15, 6002-6007.
131 Zhao, S., *et al.* (2009) Glioma-derived mutations in IDH1 dominantly inhibit IDH1 catalytic activity and induce HIF-1alpha. *Science* 324, 261-265.
132 Jin, G., *et al.* (2011) 2-hydroxyglutarate production, but not dominant negative function, is conferred by glioma-derived NADP-dependent isocitrate dehydrogenase mutations. *PloS one* 6, e16812.
133 Williams, S.C., *et al.* (2011) R132H-mutation of isocitrate dehydrogenase-1 is not sufficient for HIF-1alpha upregulation in adult glioma. *Acta neuropathologica* 121, 279-281.
134 Borger, D.R., *et al.* (2012) Frequent Mutation of Isocitrate Dehydrogenase (IDH)1 and IDH2 in Cholangiocarcinoma Identified Through Broad-Based Tumor Genotyping. *The Oncologist* 17, 72-79.
135 Figueroa, M.E., *et al.* (2010) Leukemic IDH1 and IDH2 Mutations Result in a Hypermethylation Phenotype, Disrupt TET2 Function, and Impair Hematopoietic Differentiation. *Cancer cell* 18, 553-567.
136 Sanson, M., *et al.* (2009) Isocitrate Dehydrogenase 1 Codon 132 Mutation Is an Important Prognostic Biomarker in Gliomas. *Journal of Clinical Oncology* 27, 4150-4154.
137 Xu, W., *et al.* (2011) Oncometabolite 2-Hydroxyglutarate Is a Competitive Inhibitor of α-Ketoglutarate-Dependent Dioxygenases. *Cancer cell* 19, 17-30.

138 Hewitson, K.S., *et al.* (2007) Structural and mechanistic studies on the inhibition of the hypoxia-inducible transcription factor hydroxylases by tricarboxylic acid cycle intermediates. *The Journal of biological chemistry* 282, 3293-3301.

139 Koivunen, P., *et al.* (2012) Transformation by the (R)-enantiomer of 2-hydroxyglutarate linked to EGLN activation. *Nature* 483, 484-488.

140 Losman, J.-A., *et al.* (2013) (R)-2-Hydroxyglutarate Is Sufficient to Promote Leukemogenesis and Its Effects Are Reversible. *Science* 339, 1621-1625.

141 Lu, C., *et al.* (2012) IDH mutation impairs histone demethylation and results in a block to cell differentiation. *Nature* 483, 474-478.

142 Turcan, S., *et al.* (2012) IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature* 483, 479-483.

143 Preusser, M., *et al.* (2011) Value and limitations of immunohistochemistry and gene sequencing for detection of the IDH1-R132H mutation in diffuse glioma biopsy specimens. *Journal of neuropathology and experimental neurology* 70, 715-723.

144 Jubb, H., *et al.* (2012) Structural biology and drug discovery for protein–protein interactions. *Trends in Pharmacological Sciences* 33, 241-248.

145 Amzel, L.M. (1998) Structure-based drug design. *Curr Opin Biotechnol* 9, 366-369.

146 Hung, A.W., *et al.* (2011) Route to three-dimensional fragments using diversity-oriented synthesis. *Proceedings of the National Academy of Sciences* 108, 6799-6804.

147 Bunnage, M.E. (2011) Getting pharmaceutical R&D back on target. *Nat Chem Biol* 7, 335-339.

148 Wells, J.A. and McClendon, C.L. (2007) Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* 450, 1001-1009.

149 Bernal, F., *et al.* (2010) A Stapled p53 Helix Overcomes HDMX-Mediated Suppression of p53. *Cancer cell* 18, 411-422.

150 Nooren, I.M. and Thornton, J.M. (2003) Diversity of protein-protein interactions. *EMBO J* 22, 3486-3492.

151 Wendt, M.D., *et al.* (2007) Discovery of a novel small molecule binding site of human survivin. *Bioorganic & Medicinal Chemistry Letters* 17, 3122-3129.

152 Chettiar, S.N., *et al.* (2013) Design, synthesis and biological studies of Survivin Dimerization Modulators that prolong mitotic cycle. *Bioorganic & Medicinal Chemistry Letters* 23, 5429-5433.

153 Sugase, K., *et al.* (2007) Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* 447, 1021-1025.

154 Best, J.L., *et al.* (2004) Identification of small-molecule antagonists that inhibit an activator:coactivator interaction. *Proceedings of the National Academy of Sciences of the United States of America* 101, 17622-17627.

155 Conte, L.L., *et al.* (1999) The atomic structure of protein-protein recognition sites. *Journal of Molecular Biology* 285, 2177-2198.

156 Jones, S. and Thornton, J.M. (1996) Principles of protein-protein interactions. *Proceedings of the National Academy of Sciences* 93, 13-20.

157 Bogan, A.A. and Thorn, K.S. (1998) Anatomy of hot spots in protein interfaces. *Journal of Molecular Biology* 280, 1-9.

158 Basse, M.J., *et al.* (2013) 2P2Idb: a structural database dedicated to orthosteric modulation of protein–protein interactions. *Nucleic Acids Research* 41, D824-D827.

159 Bourgeas, R., *et al.* (2010) Atomic Analysis of Protein-Protein Interfaces with Known Inhibitors: The 2P2I Database. *PloS one* 5, e9598.

160 Erlanson, D.A., *et al.* (2000) Site-directed ligand discovery. *Proceedings of the National Academy of Sciences* 97, 9367-9372.

161 Higueruelo, A.P., *et al.* (2013) Protein–protein interactions as druggable targets: recent technological advances. *Current Opinion in Pharmacology* 13, 791-796.

162 Sievers, F., *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7.

163 Gouet, P., *et al.* (2003) ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Research* 31, 3320-3323.

164 Otwinowski, Z. and Minor, W. (1997) [20] Processing of X-ray diffraction data collected in oscillation mode. In *Methods in Enzymology* (Charles W. Carter, Jr., ed), pp. 307-326, Academic Press.

165 Terwilliger, T.C. (1994) MAD phasing: Bayesian estimates of F(A). *Acta crystallographica. Section D, Biological crystallography* 50, 11-16.

166 Terwilliger, T.C. (2000) Maximum-likelihood density modification. *Acta crystallographica. Section D, Biological crystallography* 56, 965-972.

167 Terwilliger, T.C. (2003) Automated main-chain model building by template matching and iterative fragment extension. *Acta crystallographica. Section D, Biological crystallography* 59, 38-44.

168 Terwilliger, T.C. (2003) Automated side-chain model building and sequence assignment by template matching. *Acta crystallographica. Section D, Biological crystallography* 59, 45-49.

169 Emsley, P., *et al.* (2010) Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* 66, 486-501.

170 Adams, P.D., *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography* 66, 213-221.

171 Adams, P.D., *et al.* (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallographica Section D* 58, 1948-1954.

172 Murshudov, G.N., *et al.* (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta crystallographica. Section D, Biological crystallography* 53, 240-255.

173 Schrodinger, LLC (2010) The PyMOL Molecular Graphics System, Version 1.3r1.

174 Rohle, D., *et al.* (2013) An inhibitor of mutant IDH1 delays growth and promotes differentiation of glioma cells. *Science* 340, 626-630.

175 Krissinel, E. (2010) Crystal contacts as nature's docking solutions. *Journal of Computational Chemistry* 31, 133-143.

176 Krissinel, E. and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372, 774-797.

177 Yang, B., *et al.* (2010) Molecular mechanisms of "off-on switch" of activities of human IDH1 by tumor-associated mutation R132H. *Cell research* 20, 1188-1200.

178 Edgerton, D.S., *et al.* (2009) Effects of Insulin on the Metabolic Control of Hepatic Gluconeogenesis In Vivo. *Diabetes* 58, 2766-2775.

179 Petersen, K.F., *et al.* (1998) Mechanism by which glucose and insulin inhibit net hepatic glycogenolysis in humans. *The Journal of Clinical Investigation* 101, 1203-1209.

180 Baron, A.D., *et al.* (1987) Role of hyperglucagonemia in maintenance of increased rates of hepatic glucose output in type II diabetics. *Diabetes* 36, 274-283.

181 Dunning, B.E. and Gerich, J.E. (2007) The role of alpha-cell dysregulation in fasting and postprandial hyperglycemia in type 2 diabetes and therapeutic implications. *Endocrine reviews* 28, 253-283.

182 He, L., *et al.* (2009) Metformin and Insulin Suppress Hepatic Gluconeogenesis through Phosphorylation of CREB Binding Protein. *Cell* 137, 635-646.

183 Koo, S.-H., *et al.* (2005) The CREB coactivator TORC2 is a key regulator of fasting glucose metabolism. *Nature* 437, 1109-1111.

184 Brindle, P., *et al.* (1993) Protein-kinase-A-dependent activator in transcription factor CREB reveals new role for CREM repressers. *Nature* 364, 821-824.

185 Quinn, P.G. (1993) Distinct activation domains within cAMP response element-binding protein (CREB) mediate basal and cAMP-stimulated transcription. *The Journal of biological chemistry* 268, 16999-17009.

186 de Groot, R.P., *et al.* (1993) Multiple and cooperative phosphorylation events regulate the CREM activator function. *EMBO J* 12, 3903-3911.

187 Schumacher, M.A., *et al.* (2000) The Structure of a CREB bZIP·Somatostatin CRE Complex Reveals the Basis for Selective Dimerization and Divalent Cation-enhanced DNA Binding. *Journal of Biological Chemistry* 275, 35242-35247.

188 Kerppola, T. and Curran, T. (1995) Zen and the art of Fos and Jun. *Nature* 373, 199-200.

189 Foulkes, N.S., *et al.* (1991) CREM gene: Use of alternative DNA-binding domains generates multiple antagonists of cAMP-induced transcription. *Cell* 64, 739-749.

190 Gonzalez, G.A., *et al.* (1989) A cluster of phosphorylation sites on the cyclic AMP-regulated nuclear factor CREB predicted by its sequence. *Nature* 337, 749-752.

191 Hai, T.W., *et al.* (1989) Transcription factor ATF cDNA clones: an extensive family of leucine zipper proteins able to selectively form DNA-binding heterodimers. *Genes & development* 3, 2083-2090.

192 Hoeffler, J., *et al.* (1988) Cyclic AMP-responsive DNA-binding protein: structure based on a cloned placental cDNA. *Science* 242, 1430-1433.

193 Montminy, M.R., *et al.* (1986) Identification of a cyclic-AMP-responsive element within the rat somatostatin gene. *Proceedings of the National Academy of Sciences* 83, 6682-6686.

194 Short, J.M., *et al.* (1986) Characterization of the phosphoenolpyruvate carboxykinase (GTP) promoter-regulatory region. II. Identification of cAMP and

glucocorticoid regulatory domains. *The Journal of biological chemistry* 261, 9721-9726.

195 Craig, J.C., *et al.* (2001) Consensus and Variant cAMP-regulated Enhancers Have Distinct CREB-binding Properties. *Journal of Biological Chemistry* 276, 11719-11728.

196 Fink, J.S., *et al.* (1988) The CGTCA sequence motif is essential for biological activity of the vasoactive intestinal peptide gene cAMP-regulated enhancer. *Proceedings of the National Academy of Sciences* 85, 6662-6666.

197 Zhang, X., *et al.* (2005) Genome-wide analysis of cAMP-response element binding protein occupancy, phosphorylation, and target gene activation in human tissues. *Proceedings of the National Academy of Sciences of the United States of America* 102, 4459-4464.

198 Hermanson, O., *et al.* (2002) Nuclear receptor coregulators: multiple modes of modification. *Trends in endocrinology and metabolism: TEM* 13, 55-60.

199 Goodman, R.H. and Smolik, S. (2000) CBP/p300 in cell growth, transformation, and development. *Genes & development* 14, 1553-1577.

200 Kwok, R.P.S., *et al.* (1994) Nuclear protein CBP is a coactivator for the transcription factor CREB. *Nature* 370, 223-226.

201 Yao, T.P., *et al.* (1998) Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. *Cell* 93, 361-372.

202 Chrivia, J.C., *et al.* (1993) Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature* 365, 855-859.

203 Ponting, C.P., *et al.* (1996) ZZ and TAZ: new putative zinc fingers in dystrophin and other proteins. *Trends in biochemical sciences* 21, 11-13.

204 Thakur, J.K., *et al.* (2013) Molecular recognition by the KIX domain and its role in gene regulation. *Nucleic Acids Research*.

205 Deng, Z., *et al.* (2003) The CBP Bromodomain and Nucleosome Targeting Are Required for Zta-Directed Nucleosome Acetylation and Transcription Activation. *Molecular and Cellular Biology* 23, 2633-2644.

206 Bedford, D.C., *et al.* (2010) Target gene context influences the transcriptional requirement for the KAT3 family of CBP and p300 histone acetyltransferases. *Epigenetics : official journal of the DNA Methylation Society* 5, 9-15.

207 Conkright, M.D., *et al.* (2003) TORCs: Transducers of Regulated CREB Activity. *Molecular cell* 12, 413-423.

208 Schumacher, M.A., *et al.* (2000) The structure of a CREB bZIP.somatostatin CRE complex reveals the basis for selective dimerization and divalent cation-enhanced DNA binding. *The Journal of biological chemistry* 275, 35242-35247.

209 Parker, D., *et al.* (1996) Phosphorylation of CREB at Ser-133 induces complex formation with CREB- binding protein via a direct mechanism. *Molecular and Cellular Biology* 16, 694-703.

210 Yoon, J.C., *et al.* (2001) Control of hepatic gluconeogenesis through the transcriptional coactivator PGC-1. *Nature* 413, 131-138.

211 Zhou, X.Y., *et al.* (2004) Insulin regulation of hepatic gluconeogenesis through phosphorylation of CREB-binding protein. *Nat Med* 10, 633-637.

212 Dentin, R., *et al.* (2007) Insulin modulates gluconeogenesis by inhibition of the coactivator TORC2. *Nature* 449, 366-369.

213 Dentin, R., *et al.* (2008) Hepatic glucose sensing via the CREB coactivator CRTC2. *Science* 319, 1402-1405.

214 Le Lay, J., *et al.* (2009) CRTC2 (TORC2) contributes to the transcriptional response to fasting in the liver but is not required for the maintenance of glucose homeostasis. *Cell metabolism* 10, 55-62.

215 Rothwell, N.J., *et al.* (1982) Diet-induced thermogenesis. *Pharmacology & Therapeutics* 17, 251-268.

216 Thomas, S.A. and Palmiter, R.D. (1997) Thermoregulatory and metabolic phenotypes of mice lacking noradrenaline and adrenaline. *Nature* 387, 94-97.

217 Hausberg, M., *et al.* (2002) Leptin Potentiates Thermogenic Sympathetic Responses to Hypothermia: A Receptor-Mediated Effect. *Diabetes* 51, 2434-2440.

218 Arner, P. (2005) Human fat cell lipolysis: Biochemistry, regulation and clinical role. *Best Practice & Research Clinical Endocrinology & Metabolism* 19, 471-482.

219 Himms-Hagen, J. (1989) Brown adipose tissue thermogenesis and obesity. *Progress in Lipid Research* 28, 67-115.

220 Holm, C. (2003) Molecular mechanisms regulating hormone-sensitive lipase and lipolysis. *Biochemical Society transactions* 31, 1120-1124.

221 Souza, S.C., *et al.* (2002) Modulation of Hormone-sensitive Lipase and Protein Kinase A-mediated Lipolysis by Perilipin A in an Adenoviral Reconstituted System. *Journal of Biological Chemistry* 277, 8267-8272.

222 Rizack, M.A. (1964) Activation of an Epinephrine-Sensitive Lipolytic Activity from Adipose Tissue by Adenosine 3',5'-Phosphate. *The Journal of biological chemistry* 239, 392-395.

223 Mottagui-Tabar, S., *et al.* (2003) Evidence for an important role of perilipin in the regulation of human adipocyte lipolysis. *Diabetologia* 46, 789-797.

224 Lusk, G. (1928) *The Elements of the Science of Nutrition*. W.B. Saunders Company.

225 le Roux, C.W. and Bloom, S.R. (2005) Peptide YY, appetite and food intake. *Proceedings of the Nutrition Society* 64, 213-216.

226 Adrian, T.E., *et al.* (1986) Peptide YY abnormalities in gastrointestinal diseases. *Gastroenterology* 90, 379-384.

227 Batterham, R.L., *et al.* (2002) Gut hormone PYY3-36 physiologically inhibits food intake. *Nature* 418, 650-654.

228 Leemhuis, J., *et al.* (2002) The Protein Kinase A Inhibitor H89 Acts on Cell Morphology by Inhibiting Rho Kinase. *Journal of Pharmacology and Experimental Therapeutics* 300, 1000-1007.

229 Slabinski, L., *et al.* (2007) XtalPred: a web server for prediction of protein crystallizability. *Bioinformatics* 23, 3403-3405.

230 Slabinski, L., *et al.* (2007) The challenge of protein structure determination—lessons from structural genomics. *Protein Science* 16, 2472-2482.

231 Bernado, P. and Svergun, D.I. (2012) Analysis of intrinsically disordered proteins by small-angle X-ray scattering. *Methods Mol Biol* 896, 107-122.

232 Jacques, D.A. and Trewhella, J. (2010) Small-angle scattering for structural biology—Expanding the frontier while avoiding the pitfalls. *Protein Science* 19, 642-657.

233 Rambo, R.P. and Tainer, J.A. (2010) Bridging the solution divide: comprehensive structural analyses of dynamic RNA, DNA, and protein assemblies by small-angle X-ray scattering. *Curr Opin Struct Biol* 20, 128-137.

234 Blanchet, C.E. and Svergun, D.I. (2013) Small-Angle X-Ray Scattering on Biological Macromolecules and Nanocomposites in Solution. *Annual Review of Physical Chemistry* 64, 37-54.

235 Mertens, H.D.T. and Svergun, D.I. (2010) Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *Journal of Structural Biology* 172, 128-141.

236 Hura, G.L*., et al.* (2009) Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). *Nat Meth* 6, 606-612.

237 Svergun, D. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography* 25, 495-503.

238 Konarev, P.V*., et al.* (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *Journal of Applied Crystallography* 36, 1277-1282.

239 Franke, D. and Svergun, D.I. (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *Journal of Applied Crystallography* 42, 342-346.

240 Volkov, V.V. and Svergun, D.I. (2003) Uniqueness of ab initio shape determination in small-angle scattering. *Journal of Applied Crystallography* 36, 860-864.

241 Radhakrishnan, I*., et al.* (1997) Solution Structure of the KIX Domain of CBP Bound to the Transactivation Domain of CREB: A Model for Activator:Coactivator Interactions. *Cell* 91, 741-752.

242 Mossessova, E. and Lima, C.D. (2000) Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Molecular cell* 5, 865-876.

Curriculum Vitae

Akunna M. Iheanacho, Ph.D.

Johns Hopkins University School of Medicine Shagamu, Ogun State, Nigeria
725 N. Wolfe Street,
Baltimore, MD 21205

EDUCATION

1/2014	Johns Hopkins University School of Medicine, Baltimore, MD Ph.D, Cellular and Molecular Physiology Adviser: Dr. Sandra Gabelli Dissertation: *Cancer-Associated Isocitrate Dehydrogenase Mutation: Structural Basis for Enzyme Inactivation*
5/2006	University of Virginia, Charlottesville, VA M.S., Chemistry Adviser: Dr. Milton Brown *Thesis: The Chemical Biology of SC-271: A Novel Microtubule Inhibitor*
5/2000	Oakwood University, Huntsville, AL B.S., Biology Major, Chemistry Minor Cum Laude

RESEARCH EXPERIENCE

8/2007-12/2013 Johns Hopkins University School of Medicine, Baltimore, MD
Graduate Research Assistant, Department of Physiology
Structural and Biophysical Characterization of Proteins

- Optimized expression and developed purification protocols for diverse set of proteins and identified mechanism of action for enzyme inhibitor.
- Displayed proficiency in biochemistry, biophysical chemistry, and microbiology.
- Advanced knowledge of molecular architecture of enzyme.
- Collaborated with interdisciplinary team.
- Presented findings at local meeting.

Phenotyping of Mouse Model of Metabolic Disease

- Pioneered use of indirect calorimetry to assess energy expenditure of mouse cohort and collaborated with an interdisciplinary team.
- Devised and implemented protocols to assess metabolic profile of mouse cohort.
- Displayed proficiency in molecular endocrinology, physiology, and cell biology.
- Customized PCR protocols for genotyping and breeding.
- Initiated and successfully competed for funding.

6/2006 - 6/2007 Johns Hopkins University School of Medicine, Baltimore, MD
Senior Research Technician, Department of Physiology
Cellular Mechanisms of Protein Transport

- Analyzed yeast orthologue of *PARK9* gene linked to Parkinson's Disease.
- Utilized biochemical assays to investigate drug sensitivity, calcium signaling, and salt tolerance in *S. cerevisiae.*

8/2002 – 5/2006 The University of Virginia, Charlottesville, VA
Graduate Research Assistant, Department of Chemistry
Biological Evaluation of Novel Cancer Therapeutics

- Devised method to quantify tubulin levels in mammalian and endothelial cell lines and constructed stable cell line to assess dose response.
- Established correlation between compound sensitivity and tubulin levels.
- Displayed proficiency in biochemistry, cancer biology, and cell biology.
- Presented findings at local and national meetings.

TEACHING EXPERIENCE

Cancer in the Under-Privileged, Indigent, or Disadvantaged Summer Fellowship:
Trained one medical student in techniques for protein expression and purification, Gabelli Laboratory, 2013.

Ongoing research in the laboratory of Dr. Gabelli:
Provided one-on-one instruction for 2 post-doctoral fellows and 2 undergraduate students in theory and techniques related to biophysical characterization of proteins, as well as techniques for protein expression and purification, 2010-2012.

Classes taught and assisted:

Grader:	Pathways and Regulation; Johns Hopkins University SOM, graded exams and input grades, 2008-2013.
Grader:	Organ Systems Physiology; Johns Hopkins University SOM, proctored and graded exams, input grades, 2008-2013.
Teaching Assistant:	Introductory Chemistry Laboratory sections I and II; University of Virginia, ran labs and demonstrated laboratory technique, held office hours, graded lab reports and input grades, 2003-2006.
Teaching Assistant:	Solution Chemistry Laboratory sections; University of Virginia, ran labs and demonstrated laboratory technique, held office hours, graded quizzes and lab reports, 2004-2005.
Teaching Assistant:	Biological Chemistry Laboratory sections I and II; University of Virginia, ran labs and demonstrated laboratory technique, wrote and graded quizzes, devised rubric for lab reports, held office hours, graded lab reports and input grades, 2002-2003.

HONORS AND AWARDS

Endocrine Society Trainee Day Travel Award	2011
NIH Research Supplement to Promote Diversity, JHUSOM	2008-2010
National Dean's List	1998-1999
Alpha Chi Honor Society	1999
Phi Eta Sigma Honor Society	1997
UNCF Full Tuition Scholarship, Oakwood University	1996-2000

PROFESSIONAL AFFILIATIONS

American Association for the Advancement of Science	2011-2013
Biophysical Society	2010-2013
Association for Women in Science	2009-2013
Endocrine Society	2008-2012
Biomedical Scholars Association	2008-2013
American Chemical Society	2004-2013

PUBLICATIONS

Iheanacho, A., Schmidt-Kittler, O., Jakonic, J., Zhou, S., Kinzler, K. W., Amzel, L. M., Vogelstein, B., Gabelli, S. "Structural Basis of Allosteric Inhibition of Mutant Isocitrate Dehydrogenase 1," *manuscript in preparation*

PRESENTATIONS

Conference contributions:

Poster: "Structural Basis of Allosteric Inhibition of Mutant Isocitrate Dehydrogenase." Research Symposium hosted by Institute for Biophysical Research, Towson, MD, September 2013.

Poster: "SC-2-71: A novel small molecule dual inhibitor of human cancer cell proliferation and angiogenesis." American Chemical Society National Meeting, Washington, DC, June 2005.

Seminars:

Presentation: "Slow Burn: Effects of High Fat Diet on Creb Binding Protein (CBP) Mutant Mice."Johns Hopkins University School of Medicine, Department of Physiology, Baltimore, MD, December 2012.

Presentation: “Control of Hepatic Gluconeogenesis by the cAMP coactivator complex.” Johns Hopkins University School of Medicine, Department of Physiology, Baltimore, MD, November 2010

Presentation: “Role of Creb Binding Protein in Regulating Hepatic Gluconeogenesis,” Johns Hopkins University School of Medicine, Department of Physiology, Baltimore, MD, September 2008.

PROFESSIONAL SERVICE

Organizer:

Physiology Journal Club; facilitated discussion of cutting-edge techniques, organized bi-weekly meetings of graduate students and post-doctoral fellows, Physiology Department, Johns Hopkins University School of Medicine, 2008-2011.

Committee Member:

Cross-Race/Cross-Gender Student-Faculty Mentoring; served as liaison between students and faculty members, addressed need for diversity training and awareness among faculty, assisted in development of interactive learning tools for faculty members, Johns Hopkins University School of Medicine, 2009-2010.

Invited Lectureship and Seminar Organization:

“Diverse Careers in Science Series”; assisted in organizing three-part seminar series for graduate students with experts from different scientific fields, Biomedical Scholars Association, Johns Hopkins University School of Medicine, 2009-2010.

“James E. K. Hildreth Lecture”; assisted in establishing annual lectureship; planning and organizing program events, Biomedical Scholars Association, Johns Hopkins University School of Medicine, 2009-2010.